CURSO DE ANÁLISIS MATEMÁTICO

L.D. KUDRIÁVTSEV

2

3 s. antes de n.e.
ARQUÍMEDES

1601 P. FERMAT 1665

1596 R.DESCARTES 1650

1643 I.NEWTON 1727

1707 L. EULER 1783

1646 G. LEIBNIZ 1716

1685 B.TAYLOR 1731

1717 J. D'ALEMBERT 1783

1736 J.LAGRANGE 1813

1749 P. LAPLACE 1827

1777 C. GAUSS 1855

1772 J. FOURIER 1837

1789 A.CAUCHY 1857

1781 B.BOLZANO 1848

1792 N.I. LOBACHEVSKI 1856

1902 M.V. OSTROGRADSKI 1862

Л. Д. КУДРЯВЦЕВ

КУРС МАТЕМАТИЧЕСКОГО АНАЛИЗА

Том 2

МОСКВА
«ВЫСШАЯ ШКОЛА»

CURSO DE ANÁLISIS MATEMÁTICO

L. D. KUDRIÁVTSEV

2

EDITORIAL MIR MOSCÚ

Traducido del ruso por el ingeniero K. P. Medkov

Impreso en la URSS

На испанском языке

© Издательство «Высшая школа», 1981
© Traducción al español. Editorial Mir. 1984

Este libro es la segunda parte del Curso de análisis matemático en dos tomos. En el libro vienen expuestas las cuestiones, estudiadas corrientemente por los estudiantes de segundo año. La numeración de los capítulos, párrafos y figuras en este tomo continúa la numeración correspondiente del primer tomo.

El quinto capítulo con el que empieza este tomo está dedicado al cálculo diferencial de las funciones de varias variables y representa, en esencia, la continuación inmediata del segundo capítulo del primer tomo. Los capítulos ulteriores contienen la exposición del cálculo integral de las funciones de varias variables, la teoría de las series y de la integral de Fourier. La transformación de Fourier se expone al principio en la forma clásica, y luego se dan sus generalizaciones para el espacio L_2 y para las funciones generalizadas. El tomo se termina por un "Complemento", cuya parte principal concierne a los métodos numéricos destinados para calcular los valores aproximados de las funciones, las soluciones aproximadas de ecuaciones y a los cálculos aproximados de las integrales.

Indice

CAPÍTULO QUINTO

Cálculo diferencial de la función de varias variables (continuación)

§ 39. Fórmula de Taylor y serie de Taylor para las funciones de varias variables 11

39.1. Fórmula de Taylor para las funciones de varias variables 11
39.2. Fórmula de incrementos finitos para las funciones de varias variables 19
39.3. Sobre la estimación del término residual de la fórmula de Taylor en todo el dominio de definición de la función 20
39.4. Convergencia uniforme según el parámetro de una familia de funciones 23
39.5. Observaciones acerca de las series de Taylor para las funciones de varias variables 26

§ 40. Extremos de las funciones de varias variables 26
40.1. Condiciones necesarias de un extremo 26
40.2. Condiciones suficientes de un extremo estricto 29
40.3. Observaciones sobre los extremos en los conjuntos 35

§ 41. Funciones implícitas 35
41.1. Funciones implícitas definidas por una ecuación 35
41.2. Productos de los conjuntos 41
41.3. Funciones implícitas definidas por un sistema de ecuaciones 42
41.4. Aplicaciones 52
41.5. Aplicaciones vectoriales 60
41.6. Aplicaciones lineales 61
41.7. Aplicaciones derivables 67
41.8. Aplicaciones con un jacobiano distinto de cero. Principio de conservación de la región 74
41.9. Funciones implícitas definidas por una ecuación en la que se trastornan las condiciones de unicidad. Puntos singulares de las curvas planas 77
41.10. Cambio de variables 87

§ 42. Dependencia de las funciones 90
42.1. Concepto de dependencia de las funciones. Condición necesaria para la dependencia de las funciones 90
42.2. Condiciones suficientes para la dependencia de las funciones 92

§ 43. Extremo condicionado 97
43.1. Concepto de extremo condicionado 97
43.2. Método de los multiplicadores de Lagrange para buscar los puntos de extremo condicionado 101
43.3*. Interpretación geométrica del método de Lagrange 104
43.4*. Puntos estacionarios de la función de Lagrange 106
43.5. Condiciones suficientes para los puntos de extremo condicionado 111

CAPÍTULO SEXTO

Cálculo integral de las funciones de varias variables

§ 44. Integrales múltiples 116
44.1. Concepto de volumen en un espacio n-dimensional (medida de Jordan). Conjuntos medibles 116
44.2. Conjuntos de medida cero 130

44.3. Definición de la integral múltiple 134
44.4. Existencia de la integral 140
44.5* Sobre la integrabilidad de las funciones discontinuas 145
44.6. Propiedades de la integral múltiple 148
44.7* Criterios de integrabilidad de las funciones de Riemann y Darboux y los corolarios 152

§ 45. Reducción de la integral múltiple a una reiterada 160
45.1. Reducción de la integral doble a una reiterada 160
45.2. Generalización para el caso n-dimensional 167
45.3* Desigualdad integral generalizada de Minkowski 171

§ 46. Cambio de variables en una integral múltiple 173
46.1. Sentido geométrico del módulo de jacobiano en el caso bidimensional 173
46.2. Cambio de variables en una integral doble 181
46.3. Coordenadas curvilíneas 187
46.4. Cambio de variables en una integral n-múltiple 190

§ 47. Integrales curvilíneas 192
47.1. Integrales curvilíneas de primera especie 192
47.2. Integrales curvilíneas de segunda especie 195
47.3. Ampliación de la clase de transformaciones admisibles del parámetro de una curva 199
47.4. Integrales curvilíneas a lo largo de las curvas suaves a trozos 201
47.5. Fórmula de Green 202
47.6. Cálculo de las áreas mediante integrales curvilíneas 207
47.7. Significado geométrico del signo del jacobiano de la aplicación de una región plana 208
47.8. Condiciones de independencia de una integral curvilínea del camino de integración 212

§ 48. Integrales múltiples impropias 223
48.1. Definiciones fundamentales 223
48.2. Integrales impropias de las funciones no negativas 225
48.3. Integrales impropias de las funciones que cambian de signo 230

§ 49. Algunas aplicaciones geométricas y físicas de las integrales múltiples 233
49.1. Cálculo de áreas y de volúmenes 233
49.2. Aplicaciones físicas de las integrales múltiples 235

§ 50. Elementos de la teoría de superficies 237
50.1. Concepto de superficie 237
50.2*. Aplicaciones equivalentes. Superficies dadas en forma paramétrica 239
50.3. Superficies dadas implícitamente 244
50.4. Plano tangente y normal a la superficie 245
50.5. Primera forma cuadrática de una superficie 251
50.6. Curvas en una superficie. Cálculo de sus longitudes y de ángulos entre ellas 253
50.7. Area de una superficie 254
50.8. Orientación de la superficie suave 257
50.9. Pegamiento de las superficies 259
50.10. Superficies orientables y no orientables 262
50.11. Segundo enfoque del concepto de orientación de una superficie 263

§ 51. Integrales de superficie 267
51.1. Definiciones y propiedades de las integrales de superficie 267

51.2. Integrales de superficie como límites de las sumas integrales 272
51.3. Integrales de superficie extendidas a las superficies suaves a trozos 273

§ 52. Campos escalares y vectoriales 276
52.1. Definiciones 276
52.2. Sobre la invariación de los conceptos de gradiente, divergencia y rotor 281
52.3. Fórmula de Ostrogradski — Gauss. Definición geométrica de la divergencia 285
52.4. Fórmula de Stokes. Definición geométrica del rotor 290
52.5. Campos vectoriales solenoidales 295
52.6. Campos vectoriales potenciales 297
§ 53. Integrales propias dependientes de un parámetro 301
53.1. Definición de las integrales dependientes de un parámetro; su continuidad e integrabilidad según el parámetro 301
53.2. Derivación de las integrales dependientes de un parámetro 304

§ 54. Integrales impropias dependientes de un parámetro 307
54.1. Definiciones fundamentales. Convergencia uniforme de las integrales dependientes de un parámetro 307
54.2*. Criterio de la convergencia uniforme de las integrales 312
54.3. Propiedades de las integrales impropias dependientes de un parámetro 314
54.4. Aplicación de la teoría de integrales dependientes de un parámetro al cálculo de las integrales definidas. 320
54.5. Integrales de Euler 325
54.6. Funciones de valores complejos de un argumento real 330
54.7*. Comportamiento asintótico de la función gamma 332
54.8*. Series asintóticas 338
54.9*. Desarrollo asintótico de la función gamma incompleta 341
54.10. Observaciones sobre las integrales múltiples dependientes de un parámetro 344

CAPÍTULO VII

Series de Fourier. Integral de Fourier

§ 55. Series trigonométricas de Fourier 346
55.1. Definición de la serie de Fourier. Planteamiento de los problemas fundamentales 346
55.2. Coeficientes de Fourier que tienden a cero 351
55.3. Integral de Dirichlet. Principio de localización 355
55.4. Convergencia de la serie de Fourier en un punto 360
55.5*. Convergencia de las series de Fourier para las funciones que satisfacen la condición de Hölder 368
55.6. Sumación de las series de Fourier por el método de medias aritméticas 371
55.7. Aproximación de las funciones continuas por medio de los polinomios 376
55.8. Completitud del sistema trigonométrico y del sistema de potencias enteras no negativas de x en un espacio de funciones continuas 378
55.9. Propiedad mínimal de los coeficientes de Fourier. Desigualdad de Bessel e igualdad de Parseval 380
55.10. Carácter de convergencia de las series de Fourier. Derivación de las series de Fourier término a término 384
55.11. Integración de las series de Fourier término a término 389
55.12. Series de Fourier para el caso de un intervalo arbitrario. Notación compleja de las series de Fourier 391

§ 56. Integral de Fourier y transformación de Fourier 393
56.1. Representación de la función en forma de la integral de Fourier 393

56.2.	Diferentes formas de notación de la fórmula de Fourier	398
56.3.	Valor principal de una integral	399
56.4.	Notación compleja de la integral de Fourier	400
56.5.	Transformación de Fourier	400
56.6.	Integrales de Laplace	403
56.7.	Propiedades de las transformaciones de Fourier de las funciones absolutamente integrables	404
56.8.	Transformación de Fourier de las derivadas	407
56.9.	Convolución y transformación de Fourier	408
56.10.	Derivada de la transformación de Fourier de una función	411
§ 57.	Espacios funcionales	413
57.1.	Espacios métricos	413
57.2.	Espacios lineales	422
57.3.	Espacios normalizados y seminormalizados	426
57.4.	Ejemplos de espacios normalizados y seminormalizados	427
57.5.	Propiedades de los espacios seminormalizados	436
57.6.	Propiedades de los espacios normalizados	439
57.7.	Espacios lineales provistos de producto escalar	445
57.8.	Ejemplos de espacios lineales provistos de producto escalar	447
57.9.	Propiedades de los espacios lineales provistos de producto escalar. Espacios de Hilbert	449
57.10.	Espacio L_2	453
§ 58.	Bases ortonormalizadas y desarrollos según ellas	468
58.1.	Sistemas ortonormalizados	468
58.2.	Ortogonalización	472
58.3.	Sistemas completos. Completitud del sistema trigonométrico y del sistema de los polinomios de Legendre	474
58.4.	Series de Fourier	477
58.5.	Existencia de la base en los espacios separables de Hilbert. Isomorfismo de los espacios separables de Hilbert	486
58.6.	Desarrollo de las funciones de cuadrado integrable en serie de Fourier	491
58.7*.	Transformación de Fourier de las funciones integrables en cuadrado. Teorema de Plancherel	496
§ 59.	Funciones generalizadas	505
59.1.	Razonamientos generales	505
59.2.	Espacios lineales con convergencia. Funcionales. Espacios conjugados	511
59.3.	Definición de las funciones generalizadas. Espacios D y D'	514
59.4.	Derivación de las funciones generalizadas	520
59.5.	Espacio de las funciones principales S y espacio de las funciones generalizadas S'	524
59.6.	Transformación de Fourier en el espacio S	526
59.7.	Transformación de Fourier de las funciones generalizadas	529

Complemento

§ 60.	Algunos problemas de los cálculos aproximados	537
60.1.	Aplicación de la fórmula de Taylor para el cálculo aproximado de los valores de funciones e integrales	537
60.2.	Resolución de las ecuaciones	541
60.3.	Interpolación de las funciones	548
60.4.	Fórmulas de cuadratura	550
60.5.	Error de las fórmulas de cuadratura	553

60.6. Cálculo aproximado de las derivadas 559

§ 61. Partición del conjunto en clases de elementos equivalentes 561

§ 62. Límite según un filtro 562

62.1. Espacios topológicos 563
62.2. Filtros 564
62.3. Límite de un filtro 568
62.4. Límite de la aplicación según un filtro 569

Indice de nombres 571

Indice alfabético de materias 572

CAPÍTULO QUINTO

CÁLCULO DIFERENCIAL DE LA FUNCIÓN DE VARIAS VARIABLES (CONTINUACIÓN)

§ 39. FÓRMULA DE TAYLOR Y SERIE DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

39.1. FÓRMULA DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

Si una función de varias variables tiene un número suficiente de derivadas continuas en un entorno de cierto punto, dicha función puede ser representada dentro del entorno citado (al igual que se hizo en el caso de la función de una sola variable) en forma de una suma de cierto polinomio y un resto que es, en determinado sentido, "pequeño".

Teorema 1. *Supongamos que la función $z = f(x, y)$ está definida y es continua, lo mismo que todas sus derivadas parciales hasta el orden m inclusive ($m \geqslant 1$), en un δ-entorno*[*] *del punto (x_0, y_0). Entonces, para cualesquiera Δx y Δy, que satisfacen la condición $\rho = \sqrt{\Delta x^2 + \Delta y^2} < \delta$, existe tal $\theta = \theta(\Delta x, \Delta y)$, $0 < \theta < 1$, que resulta ser válida la fórmula*

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial y}\Delta y +$$

$$+ \frac{1}{2!}\left[\frac{\partial^2 f(x_0, y_0)}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f(x_0, y_0)}{\partial y^2}\Delta y^2\right] +$$

$$+ \frac{1}{3!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[3]} f(x_0, y_0) + \ldots$$

$$\ldots + \frac{1}{(m-1)!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m-1]} f(x_0, y_0) + r_{m-1}(\Delta x, \Delta y)$$

[*] En algunas obras se emplea el término "δ-vecindad". (*N. del Tr.*)

o, en la forma más breve,

$$\Delta z = \sum_{k=1}^{m-1} \frac{1}{k!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0) + r_{m-1}(\Delta x, \Delta y), \tag{39.1}$$

donde

$$r_{m-1}(\Delta x, \Delta y) = \frac{1}{m!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m]} f(x_0 + \theta\Delta x, y_0 + \theta\Delta y). \tag{39.2}$$

La fórmula (39.1) se denomina *fórmula de Taylor* (de orden $m - 1$) para la función f y la función $r_{m-1}(\Delta x, \Delta y)$, *su término residual*, mientras que la inscripción (39.2) de éste lleva el nombre de término residual de la fórmula de Taylor en la *forma de Lagrange*.

Cuando $m = 1$, en (39.1) requiere explicaciones el sentido del primer término del segundo miembro, puesto que en este caso el supraíndice de sumación es igual a cero. En el caso dado se supone, según la definición, que el término citado es igual a cero, es decir, la fórmula (39.1) adopta la forma

$$\Delta z = r_0(\Delta x, \Delta y).$$

En adelante, siempre cuando se encuentre una expresión escrita con ayuda del símbolo $\sum$, en la que el valor del supraíndice de sumación sea inferior al valor del subíndice, convendremos en considerar también que esta expresión es igual a cero.

DEMOSTRACIÓN. Supongamos que Δx y Δy están fijados de una manera tal que $\rho = \sqrt{\Delta x^2 + \Delta y^2} < \delta$, entonces todos los puntos del tipo $(x_0 + t\Delta x, y_0 + t\Delta y)$, donde $0 \leqslant t \leqslant 1$, se disponen en un segmento que une los puntos (x_0, y_0) y $(x_0 + \Delta x, y_0 + \Delta y)$, por lo cual todos ellos pertenecen al δ-entorno del punto (x_0, y_0). Por esta razón tiene sentido la composición de funciones

$$z = f(x, y)$$

y

$$x = x_0 + t\Delta x, \quad y = y_0 + t\Delta y, \quad 0 \leqslant t \leqslant 1,$$

es decir, la función compuesta

$$F(t) = f(x_0 + t\Delta x, y_0 + t\Delta y), \quad 0 \leqslant t \leqslant 1. \tag{39.3}$$

Es evidente que

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = F(1) - F(0). \tag{39.4}$$

Como la función f tiene en el δ-entorno del punto (x_0, y_0) m derivadas parciales continuas, de acuerdo con el teorema sobre las derivadas de una función compuesta (véase el p. 20.3), la función F tendrá también en el segmento $[0, 1]$ m derivadas continuas, por lo cual para F será válida la fórmula de Taylor de orden $m - 1$ con el término residual en la forma de Lagrange:

$$F(t) - F(0) = F'(0)t + \frac{F''(0)}{2!}t^2 + \ldots + \frac{F^{(m-1)}(0)}{(m-1)!}t^{m-1} + \frac{F^{(m)}(\theta t)}{m!}t^m,$$

$$0 < \theta < 1, \tag{39.5}$$

y en el entorno considerado del punto (x_0, y_0) la función (39.3) puede derivarse m veces conforme a la regla para derivar funciones compuestas (véase la observación 2 en p. 20.4), con la particularidad de que los valores de las derivadas parciales mixtas que se obtienen *no dependen* del orden en que se lleva a cabo la derivación (véase el p. 21.1).

Al expresar las derivadas $F^{(k)}(t)$ en términos de las derivadas de la función $f(x, y)$ y al poner en la fórmula (39.5) $t = 1$ (véase (39.4)), obtendremos la función requerida de Taylor correspondiente a la función $f(x, y)$. En efecto, de (39.3) se deduce que

$$F'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dx}{dt} = \\ = \frac{\partial f(x_0 + t\Delta x, y_0 + t\Delta y)}{\partial x}\Delta x + \frac{\partial f(x_0 + t\Delta x, y_0 + t\Delta y)}{\partial y}\Delta y.$$

De aquí, al omitir para abreviar las designaciones de los argumentos, obtendremos para $F''(t)$

$$F''(t) = \frac{d}{dt}\left(\frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y\right) = \frac{\partial^2 f}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}\Delta y^2.$$

En general, por inducción es fácil establecer que

$$F^{(k)}(t) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0 + t\Delta x, y_0 + t\Delta y),$$

$$k = 1, 2, \ldots, m. \quad (39.6)$$

Al poner en las fórmulas (39.6) $t = 0$ para $k = 1, 2, \ldots, m - 1$, tendremos:

$$F'(0) = \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial y}\Delta y;$$

$$F''(0) = \frac{\partial^2 f(x_0, y_0)}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f(x_0, y_0)}{\partial y^2}\Delta y^2,$$

y, en general,

$$F^{(k)}(0) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0), \quad k = 1, 2, \ldots, m - 1. \quad (39.7)$$

Cuando $k = m$, al sustituir t por θt,

$$F^{(m)}(\theta t) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m]} f(x_0 + \theta t \Delta x, y_0 + \theta t \Delta y). \quad (39.8)$$

Sustituyamos ahora (39.7) y (39.8) en (39.5) y pongamos $t = 1$; en este caso, en virtud de la relación (39.4)

$$\Delta z = F(1) - F(0) = \sum_{k=1}^{m-1} \frac{F^{(k)}(0)}{k!} + \frac{F^{(m)}(0)}{m!} =$$

$$= \sum_{k=1}^{m-1} \frac{1}{k!} \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[k]} f(x_0, y_0) +$$

$$+ \frac{1}{m!} \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[m]} f(x_0 + \theta \Delta x, y_0 + \theta \Delta y), \quad 0 < \theta < 1. \qquad \square$$

Corolario. *En vista de las suposiciones del teorema 1 es lícita la fórmula*

$$\Delta z = \sum_{k=1}^{m} \frac{1}{k!} \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[k]} f(x_0, y_0) + r_m(\Delta x, \Delta y), \tag{39.9}$$

con la particularidad de que el término residual $r_m(\Delta x, \Delta y)$ puede ser escrito en cualquiera de las siguientes formas:

$$r_m(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k}, \tag{39.10}$$

donde

$$\lim_{\rho \to 0} \varepsilon_k(\Delta x, \Delta y) = 0, \quad k = 0, 1, \ldots, m, \quad \rho = \sqrt{\Delta x^2 + \Delta y^2},$$

o bien

$$r_m(\Delta x, \Delta y) = \varepsilon(\Delta x, \Delta y) \rho^m, \tag{39.11}$$

donde $\lim_{\rho \to 0} \varepsilon(\Delta x, \Delta y) = 0$, *es decir,*

$$r_m(\Delta x, \Delta y) = o(\rho^m). \tag{39.12}$$

La representación del término residual de la fórmula de Taylor tal como se indica en (39.12) se denomina su inscripción en la *forma de Peano.*

DEMOSTRACIÓN. Pongamos

$$\varepsilon_k(\Delta x, \Delta y) = \frac{\partial^m f(x_0 + \theta \Delta x, y_0 + \theta \Delta y)}{\partial x^k \partial y^{m-k}} - \frac{\partial^m f(x_0, y_0)}{\partial x^k \partial y^{m-k}}. \tag{39.13}$$

Debido a la continuidad de todas las derivadas parciales de orden m

$$\lim_{\rho \to 0} \varepsilon_k(\Delta x, \Delta y) = 0$$

Al hacer uso de la expresión (39.13), transformemos el resto $r_{m-1}(\Delta x, \Delta y)$ (véase (39.2)), de la manera siguiente:

$$r_{m-1}(\Delta x, \Delta y) = \frac{1}{m!} \sum_{k=0}^{m} C_m^k \frac{\partial^m f(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{\partial x^k \partial y^{m-k}} \Delta x^k \Delta y^{m-k} =$$

$$= \frac{1}{m!} \sum_{k=0}^{m} C_m^k \frac{\partial^m f(x_0, y_0)}{\partial x^k \partial y^{m-k}} \Delta x^k \Delta y^{m-k} +$$

$$+ \frac{1}{m!} \sum_{k=0}^{m} C_m^k \varepsilon_k'(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k} =$$

$$= \frac{1}{m!} \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[m]} f(x_0, y_0) + \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k}, \qquad (39.14)$$

donde $\varepsilon_k(\Delta x, \Delta y) = \dfrac{C_m^k}{m!} \varepsilon_k'(\Delta x, \Delta y)$, y por esta razón

$$\lim_{\rho \to 0} \varepsilon_k(\Delta x, \Delta y) = 0. \qquad (39.15)$$

Sustituyendo (39.14) en (39.1), obtenemos la fórmula de Taylor (39.9) con el término residual en la forma (39.10).

Mostremos que el término residual (39.10) puede ser escrito en la forma (39.11). Con este fin pongamos

$$\varepsilon(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \left(\frac{\Delta x}{\rho} \right)^k \left(\frac{\Delta y}{\rho} \right)^{m-k} \qquad (39.16)$$

En este caso

$$r_m(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k} =$$

$$= \rho^m \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \left(\frac{\Delta x}{\rho} \right)^k \left(\frac{\Delta y}{\rho} \right)^{m-k} = \varepsilon(\Delta x, \Delta y) \rho^m,$$

y, como $\left| \dfrac{\Delta x}{\rho} \right| \leqslant 1$ y $\left| \dfrac{\Delta y}{\rho} \right| \leqslant 1$, de (39.15) se desprende que

$$\lim_{\rho \to 0} \varepsilon(\Delta x, \Delta y) = 0. \qquad \square$$

Aprovechando la noción de diferenciales de órdenes superiores, podemos atribuir a la fórmula de Taylor una forma más compacta que es idéntica en apariencia a la fórmula de Taylor para las funciones de una sola variable, escrita asimismo con el empleo de diferenciales. Efectivamente, puesto que (véase el p. 21.2)

$$d^k f(x, y) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[k]} f(x, y), \quad k = 0, 1, 2, \ldots, m,$$

entonces, suponiendo, para abreviar, $M_0 = (x_0, y_0)$ y $M = (x_0 + \Delta x, y_0 + \Delta y)$, podemos escribir la fórmula (39.9) en la forma

$$\Delta z = \sum_{k=1}^{m} \frac{1}{k!} d^k f(M_0) + r_m(M). \qquad (39.17)$$

La forma indicada de la fórmula de Taylor es más simple a consecuencia de lo cual resulta cómoda para recordarla.

He aquí algunas observaciones referentes a las demostraciones del teorema 1 y de su corolario. Ante todo, en las condiciones del teorema se ha exigido que la función f tenga derivadas continuas de orden hasta m inclusive en cierto δ-entorno del punto (x_0, y_0). Se podría exigir que en el entorno citado sean continuas *sólo* las derivadas de orden m, puesto que de la continuidad de éstas se infiere que son continuas también en el entorno dado todas las derivadas inferiores de la función en consideración, es decir, las derivadas de órdenes $k = 0, 1, \ldots, m-1$ (véase el p. 20.2).

Subrayemos que la continuidad de las derivadas parciales en el δ-entorno del punto (x_0, y_0) se ha usado, en primer lugar, para que las derivadas parciales con las que nos encontramos no dependan del orden en que se realiza la derivación (esto se ha usado tanto en la demostración de la fórmula de Taylor (39.1), como en la propia forma de notación de dicha fórmula) y, en segundo lugar, para que la función (39.3) pueda ser derivada m veces según la regla de derivación de la función compuesta. Fijemos la atención en que para $m = 1$ las derivadas mixtas están ausentes; entre tanto, para que haya posibilidad de derivar la función (39.3) una sola vez, según la regla de la función compuesta y, por consiguiente, para que sea válido el teorema 1, resulta suficiente una suposición más débil sobre la función en consideración f. A saber, en lugar de la suposición sobre la derivabilidad continua de la función f en el δ-entorno arriba mencionado del punto (x_0, y_0), es suficiente su derivabilidad en este entorno (véanse las definiciones 2 y 4 en el p. 20.2).

La continuidad de las derivadas parciales de orden m (en el *punto* (x_0, y_0)) se ha empleado también en la demostración del corolario del teorema 1: es necesaria para que las funciones $\varepsilon_k'(\Delta x, \Delta y)$, definidas mediante las fórmulas (39.13), tiendan a cero cuando $\rho \to 0$.

Recalquemos, además, que con las suposiciones admitidas en la fórmula (39.9) se ha demostrado que $r_m(\Delta x, \Delta y) = o(\rho^m)$ para $\rho \to 0$ no en el sentido del límite según cualquier dirección fijada, como podría parecer, a primera vista, proveniente de la demostración aducida, sino en el sentido más fuerte, esto es, en el sentido del límite en el punto (x_0, y_0) (¿por qué?).

La fórmula (39.1) puede ser un tanto generalizada, si no se tienen aspiraciones de que sea válida para todos los puntos $(x_0 + \Delta x_0, y_0 + \Delta y)$ del δ-entorno del punto (x_0, y_0), sino que se considera la fórmula citada sólo para Δx y Δy fijados. A saber, si la función f está definida y tiene derivadas parciales continuas de orden m en un conjunto abierto que contiene un segmento con los extremos (x_0, y_0) y $(x_0 + \Delta x, y_0 + \Delta y)$, entonces la fórmula (39.1) queda también justa, igual que su demostración. De esto se deduce que si la función f está definida en la región convexa G (véase el p. 18.2) y tiene en G derivadas parciales continuas de orden m, para cualesquiera dos puntos $(x_0, y_0) \in G$ y $(x_0 + \Delta x, y_0 + \Delta y) \in G$ es válida la fórmula de Taylor (39.1).

Ejercicio 1. Sea la función $f(x, y)$ continua, lo mismo que sus derivadas parciales hasta el orden m inclusive, en cierto entorno del punto (x_0, y_0). Demuéstrese que su polinomio de Taylor de orden m, es decir,

$$P(x, y) = \sum_{k=0}^{m} \frac{1}{k!} \left[(x - x_0) \frac{\partial}{\partial x} + (y - y_0) \frac{\partial}{\partial y} \right]^{(k)} f(x_0, y_0)$$

es un polinomio de mejor aproximación de la función $f(x, y)$ en un "entorno infinitamente pequeño del punto (x_0, y_0)". Esto significa lo siguiente: cualquiera que sea el polinomio $Q(x, y)$ de grado no superior a m (es decir, en cada término suyo la suma de exponentes de las potencias de las variables x e y no debe ser mayor que el número m) tal que

$$f(x, y) = Q(x, y) + o(\rho^n), \quad n \geqslant m, \quad \text{cuando} \quad \rho \to 0,$$

donde

$$\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

coincide con el polinomio citado de Taylor $P(x, y)$ de la función $f(x, y)$.

Todo lo dicho se extiende también al caso de una función de cualquier número de variables.

Teorema 1′. *Si una función de n variables $y = f(x_1, \ldots, x_n)$ está definida y es continua junto con todas sus derivadas parciales hasta el orden m, $m \geqslant 1$, inclusive en cierto δ-entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, entonces es válida la fórmula*

$$\Delta y = f(x_1^{(0)} + \Delta x_1, \ldots, x_n^{(0)} + \Delta x_n) - f(x_1^{(0)}, \ldots, x_n^{(0)}) =$$

$$= \sum_{k=1}^{m-1} \frac{1}{k!} \left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n} \right)^{(k)} f(x^{(0)}) + r_{m-1}(\Delta x), \tag{39.18}$$

donde

$$r_{m-1}(\Delta x) =$$

$$= \frac{1}{m!} \left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n} \right)^{(m)} f(x_1^{(0)} + \theta \Delta x_1, \ldots, x_n^{(0)} + \theta \Delta x_n),$$

$$0 < \theta < 1, \quad \Delta x = (\Delta x_1, \ldots, \Delta x_n), \tag{39.19}$$

y también la fórmula

$$\Delta y = \sum_{k=1}^{m} \frac{1}{k!} \left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n} \right)^{(k)} f(x^{(0)}) + r_m(\Delta x), \tag{39.20}$$

donde $r_m(\Delta x)$ puede escribirse en cada una de las siguientes formas: o bien

$$r_m(\Delta x) = \sum_{m_1 + \ldots + m_n = m} \varepsilon_{m_1 \ldots m_n}(\Delta x) \Delta x_1^{m_1} \ldots \Delta x_n^{m_n}, \tag{39.21}$$

donde $\lim_{\rho \to 0} \varepsilon_{m_1, \ldots, m_n}(\Delta x) = 0$, $\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}$, *o bien*

$$r_m(\Delta x) = \varepsilon(\Delta x) \rho^m, \quad \lim_{\rho \to 0} \varepsilon(\Delta x) = 0. \tag{39.22}$$

es decir,

$$r_m(\Delta x) = o(\rho^m), \quad \rho \to 0.$$

Por fin, en términos de las diferenciales la fórmula (39.20) *puede escribirse en la forma*

$$\Delta y = \sum_{k=1}^{m} \frac{1}{k!} d^k f(x^{(0)}) + r_m(\Delta x). \tag{39.23}$$

Ahora abrimos los paréntesis en las fórmulas (39.18) y (39.19), haciendo uso de la fórmula algebraica

$$\left(\sum_{i=1}^{n} a_i\right)^k = \sum_{k_1+\ldots+k_n=k!} \frac{k!}{k_1!k_2!\ldots k_n!} a_1^{k_1} a_2^{k_2} \ldots a_n^{k_n}.$$

Para anotar el resultado en la forma más breve introduzcamos unas designaciones nuevas. Pongamos $k = (k_1, \ldots, k_n)$, $|k| = k_1 + \ldots + k_n$, $k! = k_1! \ldots k_n!$,

$$f^{(k)} = \frac{\partial^{|k|} f}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}, \quad (x - x^{(0)})^k = (x_1 - x_1^{(0)})^{k_1} \ldots (x_n - x_n^{(0)})^{k_n};$$

$k = (k_1, \ldots, k_n)$ se llama *multiíndice.*

Introducidas las designaciones indicadas, la fórmula de Taylor (39.18) con el término residual en la forma (39.19) se escribirá así:

$$f(x) = \sum_{|k|<m} \frac{f^{(k)}(x^{(0)})}{k!}(x - x^{(0)})^k + \sum_{|k|=m} \frac{f^{(k)}(x^{(0)} + \theta(x - x^{(0)}))}{k!}(x - x^{(0)})^k.$$

Aquí, como siempre, $x = (x_1, \ldots, x_n)$, $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ y

$$x^{(0)} + \theta(x - x^{(0)}) = (x_1^{(0)} + \theta(x_1 - x_1^{(0)}), \ldots, x_n^{(0)} + \theta(x_n - x_n^{(0)})).$$

La fórmula de Taylor que acabamos de exponer para las funciones de cualquier número de variables tiene la misma forma que para las funciones de una sola variable.

A veces, particularmente en el caso de las funciones de varias variables, para las derivadas se utiliza la designación

$$D^k \overset{\text{def}}{=} \frac{\partial^{k_1+\ldots+k_n}}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}},$$

donde $k = (k_1, \ldots, k_n)$ es un multiíndice. Si se usa dicha notación, la fórmula de Taylor adoptará la forma

$$f(x) = \sum_{|k|<m} \frac{1}{k!} D^k f(x^{(0)})(x - x^{(0)})^k + \sum_{|k|=m} \frac{1}{k!} D^k f(x^{(0)} + \theta(x - x^{(0)}))(x - x^{(0)})^k, \; 0 < \theta < 1.$$

39.2. FÓRMULA DE INCREMENTOS FINITOS PARA LAS FUNCIONES DE VARIAS VARIABLES

El caso particular de la fórmula de Taylor (39.18), cuando $m = 1$, se donomina corrientemente fórmula de incrementos finitos de Lagrange para las funciones de varias variables. Debido a las observaciones que en el punto anterior siguen el teorema 1 sobre las suposiciones bajo las cuales resultan válidas las fórmulas (39.1) y (39.18), del teorema 1′ obtenemos la siguiente afirmación.

Teorema 2. *Si la función $f(x_1, \ldots, x_n)$ es derivable en todo punto de cierta región convexa $G \subset R^n$, para cualquier par de puntos $(x_1, \ldots, x_n)$ y $(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n)$ de G existe tal θ, $0 < \theta < 1$, que*

$$f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) - f(x_1, \ldots, x_n) =$$

$$= \sum_{i=1}^{n} \frac{\partial f(x_1 + \theta \Delta x_1, \ldots, x_n + \theta \Delta x_n)}{\partial x_i} \Delta x_i$$

o, en la forma más breve,

$$f(x + \Delta x) - f(x) = \sum_{i=1}^{n} \frac{\partial f(x + \theta \Delta x)}{\partial x_i} \Delta x_i, \tag{39.24}$$

donde $x = (x_1, \ldots, x_n)$, $x + \Delta x = (x_1 + \Delta x_1, \ldots, x_n + \Delta x_n)$ y

$$x + \theta \Delta x = (x_1 + \theta \Delta x_1, \ldots, x_n + \theta \Delta x_n).$$

Como ya se ha indicado, la fórmula (39.24) expresa precisamente la *fórmula de incrementos finitos de Lagrange*.

Dicha fórmula, al igual que la fórmula de Taylor en general, tiene muchas y varias aplicaciones en diferentes problemas del análisis matemático.

Prestemos nuestra atención a que el teorema 2 no es el caso particular del teorema 1, pues no es la derivabilidad continua de la función en consideración en todo punto del conjunto G lo que se exige de él, sino sólo su derivabilidad. No obstante, la demostración del teorema 2 se contiene de hecho en la demostración del teorema 1. En efecto, como ya se ha indicado en las observaciones referentes a la demostración del teorema 1 y al corolario de éste (véase el p. 39.1), para $m = 1$ la demostración del teorema 1, aducida arriba, conserva rigor también bajo las suposiciones del teorema 2, es decir, cuando sólo se supone la derivabilidad (y no derivabilidad continua) de la función f.

Demostremos la siguiente afirmación como ejemplo de la aplicación de la fórmula (39.24).

Teorema 3. *Si una función es derivable en todo punto de la región convexa G y tiene en G derivadas parciales acotadas, será uniformemente continua en la región mencionada.*

DEMOSTRACIÓN. Si

$$\left| \frac{\partial f(x)}{\partial x_i} \right| \leqslant c, \quad i = 1, 2, \ldots, n, \quad x \in G$$

(c es una constante), entonces para cualesquiera dos puntos $x' = (x_1', \ldots, x_n')$ y $x'' = (x_1'', \ldots, x_n'')$ de (39.24) se infiere que

$$|f(x'') - f(x')| \leqslant \sum_{i=1}^{n} \left| \frac{\partial f(\xi)}{\partial x_i} \right| |x_i'' - x_i'| \leqslant cn\rho(x', x'')$$

(aquí ξ es algún punto del segmento que tiene por extremos los puntos x' y x''). Por eso, si está prefijado $\varepsilon > 0$, basta tomar $\delta = \dfrac{\varepsilon}{cn}$ para que se cumpla la desigualdad

$$|f(x'') - f(x')| < \varepsilon, \tag{39.25}$$

cualesquiera que sean los puntos $x' \in G$ y $x'' \in G$ tales que $\rho(x', x'') < \delta$, lo que atestigua la continuidad uniforme de la función f en la región G. □

39.3. SOBRE LA ESTIMACIÓN DEL TÉRMINO RESIDUAL DE LA FÓRMULA DE TAYLOR EN TODO EL DOMINIO DE DEFINICIÓN DE LA FUNCIÓN

El término residual en la fórmula de Taylor depende, evidentemente, no sólo de los incrementos de los argumentos, sino también del mismo punto en cuyo entorno se examina el desarrollo de la función y que en el p. 39.1 se consideraba fijado. Ahora serán de interés para nosotros el comportamiento y la estimación del término residual en dependencia del cambio del punto mencionado. Con el fin de subrayar dicha dependencia, el término residual de orden m se designará en este párrafo por $r_m(x, \Delta x)$, donde $x = (x_1, \ldots, x_n)$ es un punto en cuyo entorno se desarrolla la función dada según la fórmula de Taylor. Como hasta ahora, $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$.

En las fórmulas (39.21) y (39.22), en lugar de $\varepsilon_{m_1 \ldots m_n}(\Delta x)$ y $\varepsilon(\Delta x)$ se escribirán $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ y $\varepsilon(x, \Delta x)$, respectivamente. En lo sucesivo nos hará falta la estimación del término residual de la fórmula de Taylor en la forma de Peano para todo el dominio en el que existe el desarrollo según la fórmula indicada.

Introduzcamos primero el concepto de continuidad de las derivadas parciales en la clausura de un conjunto abierto. Esto requiere una definición especial, puesto que en un punto límite del conjunto abierto G el concepto de derivada parcial no está definido, en el caso general, aún cuando la función quede definida en la clausura $\overline{G}$ del conjunto G (véase, por ejemplo, el punto M de la frontera de la región G en la fig. 156).

Definición 1. *Una función f, definida en el conjunto abierto $G \subset R^n$, se llama continuamente prolongable a la clausura $\overline{G}$ de éste, si existe tal función F, continua en $\overline{G}$, que $F = f$ en G.*

La función F se denomina prolongación continua de la función f (a $\overline{G}$) y se designará, para simplificar, también con el símbolo f.

En virtud de la unicidad del límite de una función es evidente que si una función definida en G tiene prolongación continua a $\overline{G}$, ésta es única.

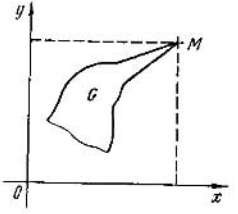


Fig. 156

Definición 2. *Una función f se denomina continuamente derivable (m veces continuamente derivable) en $\overline{G}$, si f está definida en G y todas sus derivadas parciales de primer orden (derivadas parciales hasta el orden m inclusive) son continuamente prolongables de G a $\overline{G}$.*

Ejercicios. 2. Demuéstrese que si la función f está definida en el conjunto abierto $G \subset R^n$ y tiene en éste una derivada $\frac{\partial f}{\partial x_1}$, continuamente prolongable a la clausura $\overline{G}$ del conjunto, y si, además, en cierto punto de la frontera del conjunto G existe la derivada parcial (unilateral) $\frac{\partial f}{\partial x}$, entonces la función coincide con la prolongación continua de la derivada parcial a dicho punto.

3. Demuéstrese que para que una función continua, definida en el conjunto abierto acotado $G \subset R^n$, sea continuamente prolongable a la clausura de éste, es necesario y suficiente que sea uniformemente continua en G. Pruébese que en el caso de un conjunto abierto no acotado la condición de continuidad uniforme de una función prolongable, siendo suficiente para la prolongación continua, no es necesaria.

4. Constrúyase un ejemplo de la función, continua y acotada en una región, que no pueda ser continuamente prolongada a la clausura de dicha región.

Volveremos ahora a la fórmula de Taylor. Supongamos que la función f es m veces continuamente derivable en la clausura $\overline{G}$ de un conjunto acotado abierto G. En este caso, de acuerdo con los resultados obtenidos en el p. 39.1, en todo punto $x \in G$ tiene lugar el desarrollo (39.20) de la función f según la fórmula de Taylor, con la particularidad de que $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ en la fórmula (39.21) y $\varepsilon(x, \Delta x)$ en la fórmula (39.22) tienden a cero, para $\rho \to 0$, uniformemente en el conjunto G (véase la definición en el p. 20.2), es decir, para cualquier $\varepsilon > 0$ existe tal $\delta = \delta(\varepsilon) > 0$ que si

$$\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}, < \delta \tag{39.26}$$

se tiene

$$|\varepsilon_{m_1 \ldots m_n}(x, \Delta x)| < \varepsilon \quad \text{y} \quad |\varepsilon(x, \Delta x)| < \varepsilon$$

para todos los puntos $x \in G$.

En el caso dado esto se deduce inmediatamente del método por cuyo intermedio se obtienen las funciones $\varepsilon_{m_1 \ldots m_n}$ y $\varepsilon(\Delta x)$. Efectivamente, por ser la clausura $\bar{G}$ del conjunto abierto G acotada y cerrada, las prolongaciones continuas de las derivadas parciales de orden m de la función dada a $\bar{G}$ son en ella uniformemente continuas, razón por la cual (véase la fórmula (39.13) para el caso de $n = 2$; en el caso general es válida la fórmula análoga), cumplida la condición (39.26), tenemos

$$|\varepsilon_{m_1 \ldots m_n}(x, \Delta x)| \leqslant \omega\left(\delta, \frac{\partial f}{\partial x_1^{m_1} \ldots \partial x_n^{m_n}}, \bar{G}\right). \tag{39.27}$$

Aquí el segundo miembro (módulo de continuidad de la derivada correspondiente) no depende del punto del conjunto G y tiende a cero cuando $\delta \to 0$. Por ello, de (39.27) se desprende que $\varepsilon_{m_1 \ldots m_n}$ tiende a cero en G de modo uniforme.

Ahora podemos estimar el infinitésimo $\varepsilon(\Delta x, \Delta y)$ en la fórmula (39.22). Para n natural arbitrario podemos representarla, por analogía con el caso en que $n = 2$ (véase 39.16), en la forma

$$\varepsilon(x, \Delta x) = \sum_{m_1 + \ldots + m_n = m} \varepsilon_{m_1 \ldots m_n}(x, \Delta x) \left(\frac{\Delta x_1}{\rho}\right)^{m_1} \ldots \left(\frac{\Delta x_n}{\rho}\right)^{m_n}.$$

De aquí tenemos:

$$|\varepsilon(x, \Delta x)| \leqslant \sum_{m_1 + \ldots + m_n = m} |\varepsilon_{m_1 \ldots m_n}(x, \Delta x)|. \tag{39.28}$$

En el segundo miembro de la desigualdad (39.28) figura cierto número fijo de sumandos; designémoslo con N. En virtud de que la función $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ tiende a cero en G uniformemente (lo que ya se ha demostrado), para cualquier $\varepsilon > 0$ prefijado existe tal $\delta = \delta(\varepsilon) > 0$, que, cumplida la condición $\rho(x, x + \Delta x) < \delta$, tenemos

$$|\varepsilon_{m_1 \ldots m_n}(x, \Delta x)| < \frac{\varepsilon}{N}, \quad m_1 + \ldots + m_n = m.$$

De aquí y también de la desigualdad (39.28) se deduce que

$$|\varepsilon(x, \Delta x)| < \varepsilon. \quad \square$$

Demos a conocer una estimación más en total del término residual de la fórmula de Taylor que se obtiene de la inscripción de éste en la forma de Lagrange (39.19).

Si la función f está definida en un conjunto abierto G y tiene en dicho conjunto derivadas parciales acotadas de orden m, es decir, si existe una constante $M > 0$ tal que

$$\left|\frac{\partial^m f(x)}{\partial x_1^{m_1} \ldots \partial x_n^{m_n}}\right| \leqslant M, \quad m_1 + \ldots + m_n = m, \quad x \in G, \tag{39.29}$$

entonces, cumplida la condición $\rho(x, x + \Delta x) < \delta$, para todo $x \in G$ se verifica la desigualdad

$$|r_{m-1}(x, \Delta x)| \leqslant \frac{Mn^m\delta^m}{m!}.$$

Esto proviene directamente de la fórmula (39.19), si los valores absolutos de cada sumando de su segundo miembro se estiman por medio de la desigualdad (39.29) y otra desigualdad evidente $|\Delta x_i| \leqslant \delta$.

39.4. CONVERGENCIA UNIFORME SEGÚN EL PARÁMETRO DE UNA FAMILIA DE FUNCIONES

En el punto anterior hemos tropezado con la noción de convergencia uniforme en el conjunto dado de una familia de funciones dependientes de cierto parámetro cuando este último tiende hacia los valores determinados. En calidad de tales funciones en nuestro caso intervenían $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ y $\varepsilon(x, \Delta x)$, donde el papel del parámetro lo desempeñaba Δx. Con este caso, en su forma más sencilla, ya chocamos antes, en el p. 20.2.

Enunciemos la definición de la convergencia uniforme de una familia de funciones en el caso general.

Definición 3. *Supongamos que X es un conjunto arbitrario, $Y \subset R^m$, $y^{(0)}$ es un punto del espacio R^m o uno de los infinitos*) ∞, $+\infty$, $-\infty$ (los dos últimos infinitos merecen ser considerados sólo en el caso cuando $m = 1$), con la particularidad de que la intersección de cualquier entorno reducido y_0 con el conjunto Y es no vacía. Supongamos luego que la función $\varphi(x)$ está definida para todo $x \in X$ y la función $f(x, y)$, para cualesquiera $x \in X$ e $y \in Y$.*

Suele decirse que $f(x, y)$ tiende uniformemente en el conjunto X hacia la función $\varphi(x)$ para $y \to y^{(0)}$ y se escribe en este caso

$$f(x, y) \underset{X}{\rightrightarrows} \varphi(x), \quad y \to y^{(0)}$$

si para cualquier $\varepsilon > 0$ existe tal entorno reducido $\mathring{U}(y^{(0)})$ del punto $y^{(0)}$ que se verifica la desigualdad

$$|f(x, y) - \varphi(x)| < \varepsilon, \tag{39.30}$$

cualesquiera que sean $x \in X$ e $y \in Y \cap \mathring{U}(y^{(0)})$.

En este caso la variable y se llama con frecuencia parámetro y la función $f(x, y)$, $y \in Y$, "familia de las funciones de x" (en el sentido de que esta función prefija las funciones de la variable x para diferentes valores fijos de $y \in Y$).

Por analogía con el caso de la convergencia uniforme de una sucesión de funciones (véase el p. 36.1), la condición de convergencia uniforme de las funciones según un parámetro puede enunciarse, empleando la noción habitual de límite de una función, de la manera siguiente.

La función $f(x, y)$ tiende uniformemente en el conjunto X hacia la función $\varphi(x)$ para $y \to y^{(0)}$ cuando, y sólo cuando,

---

*) En lo sucesivo los infinitos ∞, $+\infty$, $-\infty$ se llamarán también, para simplificar, puntos ("infinitamente alejados").

$$\lim_{y \to y^0} \sup_{x \in X} |f(x, y) - \varphi(x)| = 0. \tag{39.31}$$

Así pues, la condición $f(x, y) \overset{X}{\rightrightarrows} \varphi(x)$, $y \to y^{(0)}$, es equivalente a que la función $F(y) \overset{\text{def}}{=} \sup_{x \in X} |f(x, y) - \varphi(x)|$ tiende a cero cuando $y \to y^{(0)}$. La demostración de esta afirmación no es del todo difícil y es análoga al caso de la convergencia uniforme de una sucesión de funciones. Esta demostración queda al cargo del lector.

En el caso que se considera es justo también el análogo del criterio de Cauchy para la convergencia uniforme de las sucesiones.

Teorema 4 (criterio de Cauchy). *Para que la función $f(x, y)$, para $y \to y^{(0)}$, tienda uniformemente en el conjunto X hacia cierta función, es necesario y suficiente que para cualquier $\varepsilon > 0$ exista tal entorno reducido $\mathring{U}(y^{(0)})$ del punto $y^{(0)}$ que se verifique la desigualdad*

$$|f(x, y'') - f(x, y')| < \varepsilon, \tag{39.32}$$

cualesquiera que sean

$$y' \in \mathring{U}(y^{(0)}) \cap Y, \quad e \quad y'' \in \mathring{U}(y^{(0)}) \cap Y \quad y \quad x \in X.$$

En efecto, la necesidad en la condición (39.32) proviene fácilmente, como siempre en las situaciones semejantes, de la condición (39.30). Con el fin de demostrar la suficiencia, se debe probar que de la condición (39.32) se infiere que para cualquier $x \in X$ fijo existe $\lim_{y \to y^{(0)}} f(x, y)$ y que la función $f(x, y)$ tiende hacia este límite, para $y \to y^{(0)}$, uniformemente.

Se recomienda que el lector mismo compruebe todas estas afirmaciones.

Ejercicio 5. Demuéstrese: para que la función $f(x, y)$, $x \in X$, $y \in X$, tienda uniformemente en el conjunto X, cuando $y \to y^{(0)}$, hacia la función $\varphi(x)$, $x \in X$, es necesario y suficiente que para toda sucesión $y^{(n)} \in Y$, $y^n \neq y^{(0)}$, $n = 1, 2, \ldots$, que tiende a $y^{(0)}$, la sucesión $f(x, y^{(n)})$, $n = 1, 2, \ldots$, converja uniformemente sobre el conjunto X hacia la función $\varphi(x)$.

Ejemplos. 1. Examinemos una familia de funciones $f(x, y) = e^{-xy}$, donde $0 \leqslant x \leqslant 1$, $0 \leqslant y < +\infty$. Es evidente que

$$\lim_{y \to +\infty} f(x, y) = \begin{cases} 0, & \text{si} \quad x > 0, \\ 1, & \text{si} \quad x = 0 \end{cases}$$

(de este modo, la variable y es un parámetro, si se usa la terminología indicada anteriormente). Designaremos la función límite mediante $\varphi(x)$,

$$\varphi(x) = \begin{cases} 0, & \text{si} \quad x > 0, \\ 1, & \text{si} \quad x = 0. \end{cases} \tag{39.33}$$

Demostremos que la función $f(x, y)$ tiende hacia $\varphi(x)$, cuando $y \to +\infty$, de manera no uniforme. Para esto es suficiente probar que existe tal $\varepsilon_0 > 0$ que, cualquiera que sea el entorno $\mathring{U}(+\infty)$, se encontrarán $x \in [0, 1]$ e $y \in \mathring{U}(+\infty)$ tales que se verifique la desigualdad $|e^{-xy} - \varphi(x)| \geqslant \varepsilon_0$. Elijamos ε_0 tal que sea $0 < \varepsilon_0 < 1$, y un entorno arbitrario $\mathring{U}(+\infty)$. Entonces, cualquiera que sea $y \in \mathring{U}(+\infty)$, para él se verifica

$\lim_{x-0} e^{-xy} = 1$, y, por ende, existe tal $x \in (0, 1]$ que

$$|e^{-xy} - \varphi(x)| = |e^{-xy} - 0| > \varepsilon_0.$$

De este modo, en el caso dado las condiciones del criterio de Cauchy no se cumplen (véase el teorema 4).

Sin embargo, para todo a, $0 < a < 1$, la familia de funciones $f(x, y) = e^{-xy}$ tiende uniformemente, cuando $y \to +\infty$, hacia cero en el segmento $[a, 1]$. Comprobemos el cumplimiento de las condiciones del criterio de Cauchy en este caso (véase el teorema 4). Para cualquier $\varepsilon > 0$ existe un número $\eta_\varepsilon > 0$ tal que $e^{-a\eta_\varepsilon} < \varepsilon$ (basta tomar cualquier $\eta > \frac{|\ln \varepsilon|}{a}$); por esto para cualesquiera $y > \eta_\varepsilon$ y todos los $x \in [a, 1]$ tendremos

$$|e^{-xy} - 0| = e^{-xy} < e^{-a\eta_\varepsilon} < \varepsilon.$$

Por supuesto, la investigación de la convergencia uniforme de la familia de funciones en consideración podría ser ejecutada también aplicando el criterio (39.31). Efectivamente, al emplear la fórmula (39.33), obtendremos

$$\sup_{0 \leqslant x \leqslant 1} |e^{-xy} - \varphi(x)| \geqslant \sup_{0 < x \leqslant 1} e^{-xy} = 1,$$

por lo cual la condición (39.31) no se cumple a ciencia cierta. En cambio, si $0 < a < 1$, entonces

$$\lim_{y \to +\infty} \sup_{a \leqslant x \leqslant 1} |e^{-xy} - \varphi(x)| = \lim_{y \to +\infty} \sup_{a \leqslant x \leqslant 1} e^{-xy} = \lim_{y \to +\infty} e^{-ay} = 0.$$

De este modo,

$$e^{-xy} \underset{[0,1]}{\not\rightrightarrows} \varphi(x), \quad e^{-xy} \underset{[0,a]}{\rightrightarrows} \varphi(x), \quad y \to +\infty, \quad 0 < a < 1.$$

2. En el caso cuando Y es un conjunto de los números naturales, $Y = \{1, 2, 3, \ldots\}$, e $y^{(0)} = +\infty$, la definición aducida de la convergencia uniforme según un parámetro se convierte en la definición de convergencia uniforme de una sucesión de funciones $f_n(x) = f(x, n)$, $n = 1, 2, \ldots$, en el conjunto X.

3. Sea la función $f(x, y)$ continua en el rectángulo $Q = \{(x, y): -\infty < a \leqslant x \leqslant b < +\infty, -\infty < c \leqslant y \leqslant d < +\infty\}$ y sea $y_0 \in [c, d]$.

Designemos mediante $\omega(\delta, f)$ el módulo de continuidad de la función f en el rectángulo Q; entonces

$$|f(x, y) - f(x, y_0)| \leqslant \omega(|y - y_0|; f), \quad (x, y) \in Q. \tag{39.34}$$

El segundo miembro de esta desigualdad no depende de x, y, siendo uniforme la continuidad de la función f en el rectángulo Q, $\lim_{\delta \to 0} \omega(\delta; f) = 0$. Por esta razón, de la desigualdad (39.34) se desprende que en el segmento $[a, b]$ la función $f(x, y)$ tiende uniformemente hacia la función $f(x, y_0)$ cuando $y \to y_0$.

Ejercicio 6. Demuéstrese que si la familia de funciones $f(x, y)$, $x \in X \subset R^n$, $y \in Y \subset R^m$, es tal que para todo $y \in Y$ fijo las funciones $f(x, y)$ son continuas respecto de x sobre el conjunto X y tienden uniformemente sobre dicho conjunto hacia $\varphi(x)$ para $y \to y^{(0)}$, entonces $\varphi(x)$ es también continua sobre el conjunto X.

39.5. OBSERVACIONES ACERCA DE LAS SERIES DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

Si la función $f(x)$ está definida y es derivable un número infinito de veces en cierto δ-entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R^n$, entonces, para dicha función la fórmula de Taylor (39.20) será, evidentemente, válida para cualquier n natural, $n = 1, 2, \ldots$, y, además, $\sum_{i=1}^{n} \Delta x_i^2 < \delta^2$. Si, en este caso, la serie

$$\sum_{k=1}^{\infty} \frac{1}{k!} \left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n} \right)^{[k]} f(x^{(0)})$$

convergerá hacia $\Delta y = f(x) - f(x^{(0)})$ (véase el p. 38.2), se obtendrá la siguiente fórmula

$$\Delta y = f(x) - f(x^{(0)}) = \sum_{k=1}^{\infty} \frac{1}{k!} \left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n} \right)^{[k]} f(x^{(0)}),$$

donde $x = (x_1, \ldots, x_n)$ y $x_i - x_i^{(0)} = \Delta x_i$, $i = 1, 2, \ldots, n$. De aquí, trasladando $f(x^{(0)})$ al segundo miembro, obtenemos el desarrollo de la función en serie de potencias llamada *serie de Taylor* de la función f:

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left[\left(x_1 - x_1^{(0)} \right) \frac{\partial}{\partial x_1} + \ldots + \left(x_n - x_n^{(0)} \right) \frac{\partial}{\partial x_n} \right]^{[k]} f(x^{(0)}),$$

o, abriendo los corchetes,

$$f(x) = \sum_{|k|=0}^{\infty} \frac{1}{k!} D^k f(x^{(0)})(x - x^{(0)})^k,$$

donde $k = (k_1, \ldots, k_n)$ es un multiíndice.

Ejercicio 7. Desarróllese en serie de Taylor la función $f(x, y) = e^{x+y}$.

§ 40. EXTREMOS DE LAS FUNCIONES DE VARIAS VARIABLES

40.1. CONDICIONES NECESARIAS DE UN EXTREMO

Los problemas que se estudian en este párrafo y en algunos otros que siguen llevan un carácter analítico y sus demostraciones no se hacen más difíciles al aumentar el número de las variables. Es por eso que nuestra intención es considerar dichos problemas en el caso general n-dimensional, subrayando, si es necesario, sus peculiaridades específicas para los casos de $n = 2$ y $n = 3$.

Definición 1. *Supongamos que la función $f(x)$ está definida en el conjunto $X \subset R^n$. El punto $x^{(0)} \in X$ se llama punto de máximo estricto (de mínimo estricto), si*

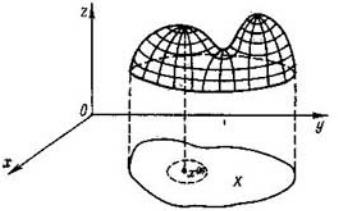


Fig. 157

existe tal entorno $U(x^{(0)})$ del punto $x^{(0)}$ que para todo $x \in U(x^{(0)}) \cap X$, $x \neq x^{(0)}$, se verifique la desigualdad $f(x) < f(x^{(0)})$ (la desigualdad $f(x) > f(x^{(0)})$ respectivamente).

De este modo, el punto de máximo estricto (de mínimo estricto) se caracteriza por que $\Delta f = f(x) - f(x^{(0)}) < 0$ ($\Delta f > 0$), cualquiera que sea $x \in U(x^{(0)}) \cap X$, $x \neq x^{(0)}$ (fig. 157).

En cambio, si para el punto $x^{(0)}$ existe un entorno $U(x^{(0)})$ tal que para todo $x \in U(x^{(0)}) \cap X$ se verifica la condición $f(x) \leqslant f(x^{(0)})$ ($f(x) \geqslant f(x^{(0)})$), entonces $x^{(0)}$ se denomina simplemente punto de máximo (punto de mínimo).

Definición 2. *Los puntos de máximo y mínimo (estrictos) de una función llevan el nombre de puntos de extremo (estricto).*

Teorema 1. *Supongamos que la función $f(x)$, $x = (x_1, x_2, \ldots, x_n)$ está definida en cierto entorno del punto $x^{(0)}$; si dicho punto es un punto de extremo de la función $f(x)$ y si en él existe cualquiera de las derivadas $\dfrac{\partial f}{\partial x_j}$ (j puede asumir uno de los valores $1, 2, \ldots, n$), esta última es nula, $\dfrac{\partial f(x^{(0)})}{\partial x_j} = 0$.*

Corolario. *Si la función $f(x)$ es derivable en el punto de extremo $x^{(0)}$, su diferencial en este punto es igual a cero, $df(x^{(0)}) = 0$.*

DEMOSTRACIÓN (del teorema y del corolario). Sea, para concretar, $j = 1$. Si $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ es el punto de extremo para la función $f(x) = f(x_1, \ldots, x_n)$, entonces $x_1^{(0)}$ es el punto de extremo para la función $f(x_1, x_2^{(0)}, \ldots, x_n^{(0)})$ de una sola variable x_1 (fig. 158), por lo cual si en este punto existe derivada $\dfrac{\partial f}{\partial x_1}$, entonces, de acuerdo con el teorema de Fermat (véase el p. 11.1), ésta es igual a cero, es decir,

$$\frac{\partial f(x^{(0)})}{\partial x_1} = \frac{df(x_1, x_2^{(0)}, \ldots, x_n^{(0)})}{dx_1}\bigg|_{x_1 = x_1^{(0)}} = 0.$$

Lo mismo ocurre en el caso de cualquier variable x_j ($j = 2, \ldots, n$).

Si la función $f(x)$ es derivable en el punto de extremo $x^{(0)}$, en este punto existen todas las derivadas $\dfrac{\partial f}{\partial x_i}$, $i = 1, 2, \ldots, n$, y, de conformidad con lo demostrado, to-

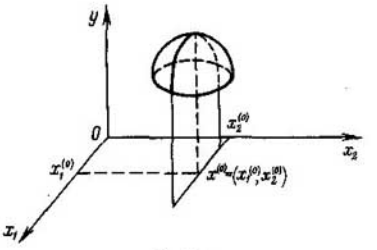


Fig. 158

das ellas son nulas, por lo cual también

$$df(x^{(0)}) = \sum_{i=1}^{n} \frac{df(x^{(0)})}{\partial x_i}\, dx_i = 0.\square$$

Ejemplos. 1. Hallemos los puntos de extremo de la función $z = x^2 + y^2$. En virtud de lo demostrado los puntos de extremo se disponen entre aquellos, para los cuales $dz = 0$. Puesto que $dz = 2x\,dx + 2y\,dy$, la condición $dz = 0$ se cumple en un único punto (0, 0). En dicho punto $z = 0$, en todos los demás $z = x^2 + y^2 > 0$. Por esta razón (0, 0) es el punto de mínimo estricto para la función $z = x^2 + y^2$ (fig. 159).

2. Investiguemos los puntos de extremo de la función $z = x^2 - y^2$. Procediendo igual que en el caso anterior, encontramos que esta vez también la condición $dz = 0$ se cumple en el punto (0, 0) y en dicho punto $z = 0$. No obstante, aquí tenemos $z > 0$ para $y = 0$ y todo $x \neq 0$, mientras que para $x = 0$ y cualquier $y \neq 0$ se tiene $z < 0$. Por esto, el punto (0, 0) no es un punto de extremo y, consecuentemente, la función $z = x^2 - y^2$ no tiene en general puntos extremos (fig. 160).

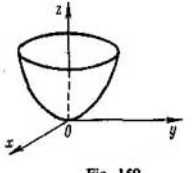


Fig. 159

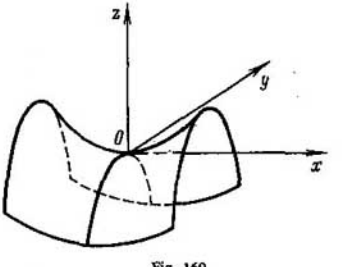


Fig. 160

40.2. CONDICIONES SUFICIENTES DE UN EXTREMO ESTRICTO

Recordemos algunas definiciones referentes al Curso del álgebra.

Definición 3. *Una forma cuadrática* $A(x) = A(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$, $a_{ji} = a_{ij}$, $i, j = 1, 2, \ldots, n$, *se denomina definida positiva (definida negativa), si* $A(x) > 0 (A(x) < 0$, *respectivamente) para cualquier punto* $x \in R^n$, $x \neq 0$.

Una forma cuadrática, que es definida positiva o definida negativa, lleva, además, el nombre de forma cuadrática de signo definido.

Definición 4. *Una forma cuadrática que asume tanto valores positivos como negativos se llama de signo indefinido.*

Lema 1. *Sea S una esfera unidad en* R^n:

$$S = \{x: x_1^2 + \ldots + x_n^2 = 1\},$$

y supongamos que $A(x)$ *es una forma cuadrática de signo definido; entonces*

$$\inf_{x \in S} |A(x)| = \mu > 0.$$

DEMOSTRACIÓN. La función $A(x)$ es un polinomio de segundo grado respecto de las variables $x_1, \ldots, x_n$, por lo cual $A(x)$ y, consecuentemente, también $|A(x)|$ son continuas en todo el espacio R^n. De aquí se desprende que la función $|A(x)|$ es continua en el compacto S. Conforme al teorema de Weierstrass, la función $|A(x)|$ alcanza en S su cota inferior, es decir, existe tal punto $x^{(0)} \in S$ que

$$\mu \stackrel{\text{def}}{=} \inf_{x \in S} |A(x)| = |A(x^{(0)})|.$$

Por definición de la forma cuadrática de signo definido, $|A(x)| > 0$ para todo punto $x \in S$, quiere decir, en particular, $\mu = |A(x^{(0)})| > 0$. □

Definición 5. *Sea f una función derivable en el punto* $x^{(0)} \in R^n$. *Si* $df(x^{(0)}) = 0$, *entonces* $x^{(0)}$ *se llama punto estacionario de la función f.*

Es evidente que el punto $x^{(0)}$ en el que la función f es derivable será estacionario, si, y sólo si,

$$\frac{df(x^{(0)})}{dx_i} = 0, \quad i = 1, 2, \ldots, n. \tag{40.1}$$

Según el corolario del teorema 1, el punto de extremo, donde la función f es derivable, es estacionario; lo recíproco, por supuesto, no es cierto en el caso general: no todo punto estacionario, en el que la función es derivable, es un punto de extremo (véase el ejemplo 2 al final del p. 40.1).

Teorema 2 (condiciones suficientes de un extremo estricto). *Supongamos que la función f está definida y tiene derivadas continuas de segundo orden en cierto entorno del punto* $x^{(0)}$. *Sea* $x^{(0)}$ *un punto estacionario de la función f; en este caso, si la forma cuadrática*

$$A(dx_1, \ldots, dx_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} dx_i, dx_j, \tag{40.2}$$

(es decir, la segunda diferencial de la función f en el punto $x^{(0)}$) es definida positiva (definida negativa), entonces $x^{(0)}$ es el punto de mínimo estricto (de máximo estricto, respectivamente); en cambio, si la forma cuadrática (40.2) *es indefinida, en el punto $x^{(0)}$ no hay extremo.*

DEMOSTRACIÓN. Sea $U(x^{(0)}, \delta_0)$ un δ_0-entorno del punto $x^{(0)}$, estacionario para la función f, en el cual la función f tiene segundas derivadas continuas. Supongamos que el punto

$$x^{(0)} + dx = (x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n)$$

pertenece a dicho entorno.

Rigiéndose por la fórmula de Taylor (véase (39.23)) y tomando en consideración las condiciones (40.1) referentes al carácter estacionario de un punto, obtenemos

$$\Delta f = f(x^{(0)} + dx) - f(x^{(0)}) = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} dx_i\, dx_j + \varepsilon(dx)\rho^2,$$

donde $dx = (dx_1, \ldots, dx_n)$, $\rho^2 = dx_1^2 + \ldots + dx_n^2$, y

$$\lim_{\rho \to 0} \varepsilon(dx) = 0, \tag{40.3}$$

o bien

$$\Delta f = \frac{\rho^2}{2}\left[\sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} \frac{dx_i}{\rho} \frac{dx_j}{\rho} + 2\varepsilon(dx) \right] =$$

$$= \frac{\rho^2}{2}\left[A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) + 2\varepsilon(dx) \right]. \tag{40.4}$$

El punto $\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$ se dispone en la esfera unidad S (es desir, en la esfera con centro en el origen de coordenadas y radio igual a 1), pues $\left(\frac{dx_1}{\rho}\right)^2 + \ldots + \left(\frac{dx_n}{\rho}\right)^2 = 1$.

Supongamos que la forma cuadrática (40.2) es de signo definido. En este caso, de acuerdo con el lema, $\inf |A| = \mu > 0$. Elijamos δ, $0 < \delta < \delta_0$, de un modo tal que sea $2|\varepsilon(dx)| < \mu$ cuando $\rho < \delta$. Entonces, para $\rho < \delta$, es decir, para $x^{(0)} + dx \in U(x^{(0)}, \delta)$ y $dx \neq 0$, toda la expresión entre corchetes en el segundo miembro de la fórmula (40.4) será del mismo signo que tiene el primer sumando $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$:

$$\operatorname{sign} \Delta f = \operatorname{sign} A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right).$$

Por eso, si la forma cuadrática (40.2) es definida positiva, tenemos $\Delta f > 0$, y si definida negativa, entonces $\Delta f < 0$, cuando $x^{(0)} + dx \in U(x^{(0)}, \delta)$. Quiere decir, en el

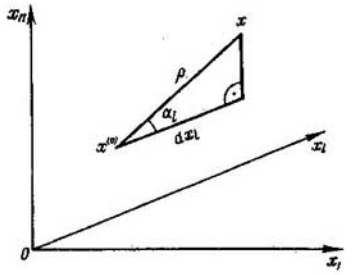


Fig. 161

primer caso $x^{(0)}$ es un punto de mínimo estricto y en el caso segundo, un punto de máximo estricto.

Supongamos ahora que la forma cuadrática (40.2) es indefinida. Esto significa que existen dos puntos $dx' = (dx_1', \ldots, dx_n')$ y $dx'' = (dx_1'', \ldots, dx_n'')$ tales que $A(dx_1', \ldots, dx_n') > 0$ y $A(dx_1'', \ldots, dx_n'') < 0$. Basándonos sobre este hecho no podemos decir de inmediato que el incremento de la función Δf cambia de signo en cualquier entorno del punto $x^{(0)}$, puesto que los puntos $x^{(0)} + dx' = (x_1^{(0)} + dx_1', \ldots, x_n^{(0)} + dx_n')$ y $x^{(0)} + dx'' = (x_1^{(0)} + dx_1'', \ldots, x_n^{(0)} + dx_n'')$ pueden, en general, incluso no pertenecer al dominio de definición de la función f. No obstante, el resultado deseado se deducirá de que la forma cuadrática $A(dx)$ conserva invariable un mismo signo o igualdad a cero en toda recta que pasa por el punto $x^{(0)}$ de la que está extraído este mismo punto, mientras que el valor $A\left(\frac{dx}{\rho}\right)$, $dx \neq 0$, no depende, en general, de la elección del punto en dicha recta.

Examinemos el punto $dx' = (dx_1', \ldots, dx_n')$. Tracemos una semirrecta que tiene su origen en el punto $x^{(0)}$ y pasa por el punto $x^{(0)} + dx'$. Para cualquier punto $x = (x_1, \ldots, x_n)$ de esta semirrecta pongamos $dx_i = x_i - x_i^{(0)}$, $i = 1, 2, \ldots, n$, y $\rho = \sqrt{\sum_{i=1}^{n} dx_i^2}$. En este caso (fig. 161)

$$\frac{dx_i}{\rho} = \cos \alpha_i, \quad i = 1, 2, \ldots, n, \tag{40.5}$$

donde $\cos \alpha_i$ son cosenos directores de la semirrecta en consideración. Por esta razón el punto

$$\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = (\cos\alpha_1, \ldots, \cos\alpha_n), \tag{40.6}$$

dispuesto, evidentemente, en la esfera unidad*) S con centro $x^{(0)}$, será el mismo para todos los puntos x de esta semirrecta, es decir, el punto (40.6) no depende de la distancia ρ entre x y $x^{(0)}$.

Por consiguiente, el valor de la forma cuadrática (40.2) en el punto (40.6), es decir, $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$ tampoco depende de ρ. De aquí, para todo punto (40.6) tenemos

$$A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = A\left(\frac{dx'_1}{\rho}, \ldots, \frac{dx'_n}{\rho}\right) = \frac{1}{\rho^2} A(dx'_1, \ldots, dx'_n) > 0.$$

Sea $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = \mu' > 0$. Elijamos $\rho_0 > 0$ de modo tal que para $\rho < \rho_0$ tenga lugar la desigualdad $2|\varepsilon(dx)| < \mu'$, lo que es posible en vista de (40.3). Entonces, para cualquier punto $x^{(0)} + dx$, que se dispone en la semirrecta (40.5) y es tal que $0 < \rho = \sqrt{\sum_{i=1}^{n} dx_i^2} < \rho_0$, la expresión entre corchetes en la fórmula (40.4) tendrá el signo del primer término y, por ende, $\Delta f > 0$. Así pues, en cualquier entorno del punto $x^{(0)}$ hay puntos para los cuales $\Delta f > 0$.

Análogamente, partiendo del valor negativo de la forma cuadrática (40.2) en el punto (dx_i'), se demuestra que en todo entorno del punto $x^{(0)}$ existen puntos, para los cuales $\Delta f < 0$. Esto es precisamente el indicio de que en el caso que se considera $x^{(0)}$ no es un punto de extremo. □

Cuando el teorema citado se aplica en la práctica, surge una pregunta: ¿cómo se establece, si es definida positiva o definida negativa la forma cuadrática (40.2)? Con este fin puede aprovecharse, por ejemplo, el así llamado *criterio de Sylvester* de la definición positiva de una forma cuadrática el cual se demuestra en el curso de álgebra. Este criterio consiste en lo siguiente.

Para que una forma cuadrática

$$A(x) = A(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \tag{40.7}$$

en la que $a_{ij} = a_{ji}$, $i, j = 1, 2, \ldots, n$, sea definida positiva, es necesario y suficiente que se verifiquen las desigualdades

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0 \ldots, \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0.$$

*) Recordemos que para los cosenos directores se verifica la igualdad $\cos^2\alpha_1 + \ldots + \cos^2\alpha_n = 1$.

Al observar que la forma cuadrática $A(x)$ es definida negativa cuando, y sólo cuando, la forma cuadrática $-A(x) = \sum_{i,j=1}^{n} (-a_{ij})\, x_i x_j$ es definida positiva, obtenemos, haciendo uso de las propiedades conocidas del determinante, el siguiente criterio para distinguir la definición negativa.

Para que la forma cuadrática (40.7) *sea definida negativa, es necesario y suficiente que se verifiquen las desigualdades*

$$a_{11} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0 \ldots, (-1)^n \begin{vmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots \\ a_{n1}\, a_{n2} \ldots a_{nn} \end{vmatrix} > 0.$$

Enunciemos ahora el teorema 2 para el caso de dos variables, expresando las condiciones que se imponen en la forma cuadrática (40.2) de una manera explícita, en términos de las segundas derivadas parciales.

Teorema 3. *Supongamos que la función $f(x, y)$ está definida y tiene derivadas parciales continuas de segundo orden en cierto entorno del punto (x_0, y_0), que es un punto estacionario para $f(x, y)$, es decir, en este punto*

$$f_x = f_y = 0. \tag{40.8}$$

Entonces, si en (x_0, y_0)

$$f_{xx} f_{yy} - f_{xy}^2 > 0, \tag{40.9}$$

es un punto de extremo estricto, a saber, de máximo estricto, si en este punto

$$f_{xx} < 0^*),$$

y de mínimo estricto, si

$$f_{xx} > 0.$$

Si en el punto (x_0, y_0)

$$f_{xx} f_{yy} - f_{xy}^2 < 0, \tag{40.10}$$

el extremo en él está ausente.

En fin, cuando

$$f_{xx} f_{yy} - f_{xy}^2 = 0 \tag{40.11}$$

en el punto (x_0, y_0), puede ocurrir que haya un extremo en él y puede ocurrir que no lo haya.

Efectivamente, si $f_{xx} \neq 0$ en el punto (x_0, y_0), la forma cuadrática (40.2) en nuestro caso puede escribirse así:

$$A(dx, dy) = f_{xx} dx^2 + 2 f_{xy} dx\, dy + f_{yy} dy^2 =$$

*) De la condición (40.9) proviene, evidentemente, que $f_{xx} \neq 0$ en el punto (x_0, y_0).

$$= \frac{1}{f_{xx}} \left[(f_{xx}dx + f_{xy}dy)^2 + (f_{xx}f_{yy} - f_{xy}^2)dy^2 \right]. \qquad (40.12)$$

Todas las derivadas parciales aquí y en adelante se refieren al punto (x_0, y_0).

Vemos inmediatamente que si se cumplen las condiciones (40.9), la expresión entre corchetes en la fórmula (40.12) es positiva para $dx^2 + dy^2 > 0$, es decir, $A(dx, dy)$ es una forma cuadrática definida, a saber, definida positiva cuando $f_{xx} > 0$ y definida negativa, cuando $f_{xx} < 0$. Por supuesto, dicha deducción proviene también del criterio de Sylvester. En el primer caso, según el teorema 2, (x_0, y_0) es el punto de mínimo estricto y en el segundo, de máximo estricto. Si, además, queda cumplida la condición (40.10), entonces para $dy = 0$, $dx \neq 0$ tenemos de (40.12): sign $A(dx, 0) =$ $=$ sign f_{xx}, y para $dx = f_{xy}$, $dy = -f_{xx}$ se obtiene sign $A(f_{xy}, -f_{xx}) = -$ sign f_{xx}, de donde se deduce que la forma cuadrática $A(dx, dy)$, es, cumplida la condición (40.10), indefinida.

Así pues, hemos investigado por completo el caso

$$f_{xx} \neq 0 \quad \text{y} \quad f_{xx}f_{yy} - f_{xy}^2 \neq 0.$$

El caso en que

$$f_{xx} = 0, \quad f_{yy} \neq 0 \quad \text{y} \quad f_{xx}f_{yy} - f_{xy}^2 \neq 0$$

se investiga de modo análogo.

Si, en cambio, $f_{xx} = f_{yy} = 0$, pero, como hasta ahora, $f_{xx}f_{yy} - f_{xy}^2 \neq 0$, entonces, evidentemente, $f_{xy} \neq 0$, por lo tanto en este caso se cumple la condición (40.10) y $A(dx, dy) = 2f_{xy}\,dx\,dy$. Vemos de inmediato que la forma cuadrática $A(dx, dy)$ es indefinida bajo las suposiciones adoptadas, pues sign $A(dx, dy) = -$ sign $A(dx, -dy)$. Por eso, con el objeto de obtener los valores de la forma cuadrática de signos contrarios, basta tomar, al principio, dx y dy de un mismo signo y después, de signos opuestos. Según el teorema 2, (x_0, y_0) no es en el caso dado un punto de extremo.

Finalmente, el caso $f_{xx} = f_{yy} = f_{xy} = 0$ no es compatible con la suposición de que $f_{xx}f_{yy} - f_{xy} \neq 0$. De este modo, hemos examinado todos los casos posibles, cuando se cumple la desigualdad $f_{xx}f_{yy} - f_{xy}^2 \neq 0$.

Para acabar con la demostración del teorema nos basta mostrar con unos ejemplos que, cuando tiene lugar la correlación (40.11), puede existir un extremo y puede no existir.

El punto (0, 0) para la función $z = x^2 + 2xy + y^2$ es estacionario y en él $z_{xx} =$ $= z_{xy} = z_{yy} = 2$, y, por tanto, se cumple la condición (40.11). Al advertir que $z =$ $= (x + y)^2$, vemos que siempre $z \geqslant 0$, con la particularidad de que $z = 0$ en la recta $x + y = 0$; por ello el punto (0, 0) es un punto de extremo, aunque sea no estricto.

Para la función $z = xy^3$ el punto (0, 0) es también estacionario y en este punto $z_{xx} = z_{yy} = z_{xy} = 0$, por lo cual la condición (40.11) queda asimismo cumplida. Sin embargo, debido a que la fórmula, que representa dicha función, contiene potencias impares de las variables x e y, la función cambia de signo en cualquier entorno de cero, a consecuencia de lo cual, (0, 0) no es un punto de extremo.

40.3. OBSERVACIONES SOBRE LOS EXTREMOS EN LOS CONJUNTOS

Sea f una función derivable en un conjunto acotado abierto G y continua en la clausura $\overline{G}$ de dicho conjunto. Se pide hallar los valores máximo y mínimo de la función f en el conjunto $\overline{G}$ (de acuerdo con el teorema 3, p. 19.5, estos valores existen). Con este objeto podemos, por ejemplo, hallar todos los puntos estacionarios de la función f en G, calcular en éstos los valores de la función y escoger, siempre que sea posible (desde el punto de vista teórico, esto es posible, por ejemplo, cuando el número de los puntos estacionarios sea finito), aquellos en los que la función asume los valores máximo y mínimo entre todos los valores en los puntos estacionarios. A continuación, conviene comparar estos valores con los que toma la función en la frontera del conjunto abierto G, hallando, por ejemplo (en el caso de poder hacerlo), los valores máximo y mínimo de la función f en la frontera de la región G. Al comparar los valores máximo y mínimo en los puntos estacionarios con los valores correspondientes máximo y mínimo en la frontera del conjunto G, podemos, evidentemente, hallar el máximo y el mínimo buscados de f en $\overline{G}$.

Cuando G es una región plana y su frontera está constituida por una curva definida por cierta representación $x = x(t)$, $y = y(t)$, $\alpha \leqslant t \leqslant \beta$, la cuestión sobre la búsqueda de los valores extremos de la función $f(x, y)$ en la frontera de G se reduce a la investigación del extremo de la función de una variable $f(x(t), y(t))$, que se realiza con ayuda de los métodos ya conocidos.

Los métodos que pueden emplearse en el caso multidimensional para buscar puntos extremos en la frontera de una región serán considerados en el § 43.

Ejercicios. 1. Hállense los extremos de la función $z = x^3 + 12xy^2 - 15x - 24y$.

2. ¿Tendrá un extremo la función $z = x^4y^2 - 3x^2y + 2x + y$ en el punto $(1, 1)$?

3. Hállense los valores máximo y mínimo de la función $z = x^2 + y^2 - 4x - 2y + 4$ en una región cerrada limitada por las líneas $x = 4$, $y = -1$, $x - y = 3$.

4. Sea $a = \text{const} > 0$, $X = \{(x, y) : |x| < a, y \in R\}$. Hállense todos los extremos de la función $z = \frac{3}{2a}x^2 + \sqrt{6(a^2 - x^2)}\cos y$ en X y todos los valores máximos y mínimos de ella en X.

5. La superficie total de un paralelepípedo rectangular es igual a $6a^2$. ¿Para qué valores de la longitud de sus aristas el volumen del paralelepípedo será máximo?

§ 41. FUNCIONES IMPLÍCITAS

41.1. FUNCIONES IMPLÍCITAS DEFINIDAS POR UNA ECUACIÓN

Aclaremos las condiciones bajo las cuales una ecuación de varias variables define una función unívoca, es decir, define una de dichas variables como función de las demás. Empezaremos nuestras consideraciones por el estudio de una ecuación que contiene dos incógnitas

$$F(x, y) = 0.$$

Si la función de dos variables $F(x, y)$ está definida en cierto subconjunto A del plano R^2_{xy}, $A \subset R^2_{xy}$, y si existe tal función de una variable $y = f(x)$, definida en el conjunto $B \subset R_x$ contenido en la proyección del conjunto A sobre el eje Ox, que para todo $x \in B$ tenga lugar $(x, f(x)) \in A$ y sea válida la identidad $F(x, f(x)) \equiv 0$, entonces f se denomina *función implícita* definida por la ecuación $F(x, y) = 0$.

Lema. *Sea $F(x, y)$ una función continua en cierto entorno rectangular*

$$U(x_0, y_0) = \{(x, y) : |x - x_0| < \xi, \quad |y - y_0| < \eta\}^*)$$

del punto (x_0, y_0) y supongamos que dicha función es, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo, estrictamente monótona respecto de y en el intervalo $(y - \eta, y_0 + \eta)$. En este caso, si

$$F(x_0, y_0) = 0,$$

existen tales entornos $U(x_0) = (x_0 - \delta, x_0 + \delta)$, del punto x_0 y $U(y_0) = (y - \varepsilon y_0 + \varepsilon)$ del punto y_0, que para todo $x \in U(x_0)$ se tiene una solución, y sólo una, $y \in U(y_0)$ de la ecuación $F(x, y) = 0$. Esta solución, que es una función de x y se designa por $y = f(x)$, es continua en el punto x_0 y

$$f(x_0) = y_0.$$

De este modo, el lema afirma, en particular, que para las suposiciones asumidas, la función implícita $y = f(x)$, definida por la ecuación $F(x, y) = 0$, existe y posee la propiedad de que las igualdades

$$F(x, y) = 0 \quad \text{e} \quad y = f(x)$$

son equivalentes a condición de que $x \in U(x_0)$, $y \in U(y_0)$.

DEMOSTRACIÓN. Por hipótesis del lema, la función $F(x, y)$ es, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo, estrictamente monótona respecto de la variable y en el intervalo $(y_0 - \eta, y_0 + \eta)$, en particular, en éste es estrictamente monótona la función $F(x_0, y)$. Supongamos, para concretar, que es estrictamente creciente. Eligamos arbitrariamente $\varepsilon > 0$, subordinado sólo a la condición $0 < \varepsilon < \eta$. Como la función $F(x_0, y)$ de la variable y es, en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$, estrictamente creciente y, por hipótesis, $F(x_0, y_0) = 0$, entonces

$$F(x_0, y_0 - \varepsilon) < 0, \quad F(x_0, y_0 + \varepsilon) > 0.$$

Mas la función de dos variables $F(x, y)$ es, conforme a la suposición, continua en el conjunto abierto $U(x_0, y_0)$ y $(x_0, y_0 - \varepsilon) \in U(x_0, y_0)$, $(x_0, y_0 + \varepsilon) \in U(x_0, y_0)$, por lo cual existe tal $\delta > 0$, $0 < \delta < \xi$, que en el δ-entorno del punto $(x_0, y_0 - \varepsilon)$ se verifica la desigualdad $F(x, y) < 0$, y en δ-entorno del punto $(x_0, y_0 + \varepsilon)$, la desigualdad $F(x, y) > 0$ (véase el lema 1 en el p. 19.3). En particular, para todo $x \in (x_0 - \delta, x_0 + \delta)$ (fig. 162) quedan válidas las desigualdades

$$F(x, y_0 - \varepsilon) < 0, \quad F(x, y_0 + \varepsilon) > 0. \tag{41.1}$$

---

*) En concordancia con las designaciones aceptadas en este libro, sería más correcto denotar el entorno del punto (x_0, y_0) mediante $U((x_0, y_0))$ en lugar de $U(x_0, y_0)$. Para simplificar las designaciones, convendremos en omitir el segundo paréntesis.

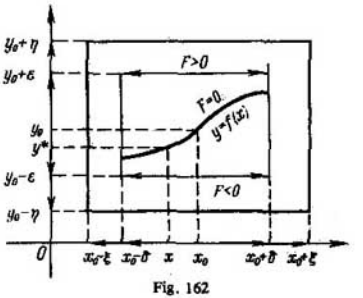


Fig. 162

Pongamos

$$U(x_0) \stackrel{\text{def}}{=} (x_0 - \delta, x_0 + \delta), \quad U(y_0) \stackrel{\text{def}}{=} (y_0 - \varepsilon, y_0 + \varepsilon).$$

Dado que, siendo fijado $x \in U(x_0)$, la función $F(x, y)$ de la variable y es continua en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$, entonces, de la condición (41.1) se deduce (de acuerdo con el teorema de Cauchy *sobre los valores intermedios de una función continua*, véase el teorema 2 en el p. 6.2) que existe tal $y^* \in U(y_0)$ (véase la fig. 162) que $F(x, y^*) = 0$. Por ser $F(x, y)$ estrictamente monótona en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$ respecto de la variable y, el valor indicado de y^* es único.

De este modo, se ha obtenido una correspondencia unívoca (función unívoca) $x \rightarrow y^*$, $x \in U(x_0)$, $y^* \in U(y_0)$ que se designará mediante f: $y^* = f(x)$.

Por definición de esta correspondencia, para cualquier $x \in U(x_0)$ e $y^* = f(x)$ tenemos

$$F(x, y^*) = 0, \quad y^* \in U(y_0)$$

con la particularidad de que el punto y^*, que posee dicha propiedad, es único. Hemos demostrado, pues, la existencia y unicidad de la función buscada f.

Luego, por hipótesis del lema, $F(x_0, y_0) = 0$, y, como $x_0 \in U(x_0)$, $y_0 \in U(y_0)$, entonces, debido a la unicidad de la función f, tenemos $y_0 = f(x_0)$.

Al final notemos que $\varepsilon > 0$ se ha fijado arbitrariamente a condición de que $\varepsilon < \eta$, y que se ha encontrado para él tal $\delta > 0$, que de $|x - x_0| < \delta$ (es decir, de la condición $x \in U(x_0)$) provenía la inclusión $f(x) \in U(y_0)$, es decir, la desigualdad $|f(x) - f(x_0)| < \varepsilon$. Esto precisamente significa la continuidad de la función f en el punto x_0. □

Las condiciones suficientes, cómodas para aplicarlas, de la resolubilidad unívoca de la ecuación $F(x, y) = 0$ en cierto entorno del punto (x_0, y_0), para el cual $F(x_0, y_0) = 0$, las proporciona el siguiente teorema.

Teorema 1. *Supongamos que la función $F(x, y)$ es continua en cierto entorno del punto (x_0, y_0) y tiene en dicho entorno una derivada parcial $F_y(x, y)$ que es continua en el punto (x_0, y_0). En este caso, si*

$$F(x_0, y_0) = 0, \quad F_y(x_0, y_0) \neq 0,$$

existen tales entornos $U(x_0)$ y $U(y_0)$ de los puntos respectivos x_0 e y_0, que para todo $x \in U(x_0)$ se tiene una y sólo una, solución $y = f(x) \in U(y_0)$ de la ecuación $F(x, y) = 0$[*]*. Esta solución es continua en todo punto de $U(x_0)$ e $y_0 = f(x_0)$.*

Si suponemos complementariamente que la función F tiene en cierto entorno del punto (x_0, y_0) una derivada parcial $F_x(x, y)$, continua en el punto (x_0, y_0), entonces la función $f(x)$ tendrá también en el punto x_0 una derivada y para ésta queda válida la fórmula

$$f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}.$$

DEMOSTRACIÓN. Debido a que la función $F(x, y)$ es continua en cierto entorno del punto (x_0, y_0) y es también continua en el mismo punto la derivada parcial $F_y(x, y)$, existe un entorno rectangular

$$U(x_0, y_0) = \{(x, y) : |x - x_0| < \xi, \quad |y - y_0| < \eta\}$$

del punto (x_0, y_0), donde la propia función $F(x, y)$ es continua y los valores de la derivada parcial $F_y(x, y)$ son del mismo signo que tiene su valor en el punto (x_0, y_0). Por esto, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo la función $\varphi(y) \overset{\text{def}}{=} F(x, y)$ es derivable en el intervalo $(y_0 - \eta, y_0 + \eta)$, mientras que su derivada $\varphi'(y) = F_y(x, y)$ conserva constante el signo. Por consiguiente, la función $\varphi(y)$ es estrictamente monótona en el intervalo indicado.

Así pues, todas las condiciones del lema para la función $F(x, y)$ vienen cumplidas en el entorno rectangular construido $U(x_0, y_0)$. Por lo tanto, existen los entornos $U(x_0) = (x_0 - \delta, x_0 + \delta)$, $U(y_0) = (y_0 - \varepsilon, y_0 + \varepsilon)$ y la única función $y = f(x)$, definida en $U(x_0)$, tales que para todo $x \in U(x_0)$ tienen lugar la inclusión $f(x) \in U(y_0)$ y la igualdad $F(x, f(x)) = 0$, con la particularidad de que la función f es continua en el punto x_0.

Dado que para todo punto (x, y), para el cual $x \in U(x_0)$, $y \in U(y_0)$, existe su entorno rectangular $U(x, y)$ contenido en otro entorno rectangular

$$U_0(x_0, y_0) = \{(x, y) : |x - x_0| < \delta, \quad |y - y_0| < \varepsilon\}$$

(fig. 163), entonces para $U(x, y)$ también se cumplen todas las condiciones del lema. Por consiguiente, siendo única la solución $f(x)$ de la ecuación $F(x, y) = 0$ en el entorno $U_0(x_0, y_0)$, la función $y = f(x)$ es continua en todo punto $x \in U(x_0)$, de conformidad con el mismo lema.

Demostremos ahora la última afirmación del teorema. En vista de la continuidad de las derivadas parciales F_x y F_y en el punto (x_0, y_0), la función F es derivable en este punto:

$$F(x_0 + \Delta x, y_0 + \Delta y) - F(x_0, y_0) = \\ = F_x(x_0, y_0)\,\Delta x + F_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y, \quad (41.2)$$

[*] En este caso suele decirse también que la ecuación $F(x, y) = 0$ es resoluble univocamente en el entorno $U(x_0, y_0) = \{(x, y) : x \in U(x_0), y \in U(y_0)\}$ del punto (x_0, y_0).

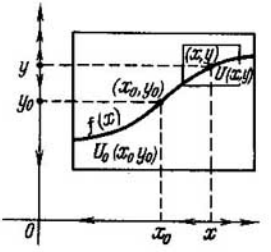


Fig. 163

donde

$$\lim_{\rho \to 0} \varepsilon_1 = \lim_{\rho \to 0} \varepsilon_2 = 0, \quad \rho = \sqrt{\Delta x^2 + \Delta y^2}.$$

Tomemos en la fórmula (41.2)

$$x_0 + \Delta x \in U(x_0), \quad \Delta y = f(x_0 + \Delta x) - f(x_0).$$

Entonces, en vista de la condición $F(x, f(x)) \equiv 0$, obtenemos

$$F(x_0 + \Delta x, y_0 + \Delta y) \equiv F(x_0 + \Delta x, f(x_0 + \Delta x)) = 0,$$

y, como $F(x_0, y_0) = 0$, de (41.2) tenemos

$$F_x(x_0, y_0)\,\Delta x + F_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y = 0.$$

De aquí

$$\frac{\Delta y}{\Delta x} = -\frac{F_x(x_0, y_0) + \varepsilon_1}{F_y(x_0, y_0) + \varepsilon_2}. \tag{41.3}$$

Sea ahora $\Delta x \to 0$; entonces, siendo la función f continua, $\Delta y \to 0$, y esto significa que cuando $\Delta x \to 0$, se tiene $\rho = \sqrt{\Delta x^2 + \Delta y^2} \to 0$, de donde se desprende que en la fórmula (41.3) $\lim\limits_{\Delta x \to 0} \varepsilon_1 = \lim\limits_{\Delta x \to 0} \varepsilon_2 = 0$. Es por esto que cuando $\Delta x \to 0$, el límite del segundo miembro de la igualdad (41.3) existe y es igual a $-\dfrac{F_x(x_0, y_0)}{F_y(x_0, y_0)}$ (recordemos que $F_y(x_0, y_0) \neq 0$), por consiguiente, cuando $\Delta x \to 0$, existe también el límite del primer miembro, es decir, existe la derivada

$$f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}. \square \tag{41.4}$$

OBSERVACIÓN. Si las funciones F_x y F_y son continuas en el entorno $U_0(x_0, y_0)$ del punto (x_0, y_0), la derivada f' es continua en el intervalo $U(x_0)$. Efectivamente, al

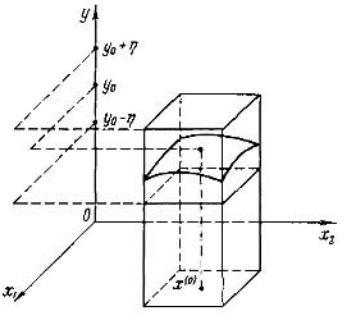


Fig. 164

aplicar la fórmula (41.4) a un punto arbitrario $x \in U(x_0)$, obtenemos

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))},$$

de donde proviene, según el teorema sobre la composición de funciones continuas, la continuidad de la función $f'(x)$ en $U(x_0)$.

Análogamente se introduce la noción de función implícita definida por la ecuación

$$F(x_1, \ldots, x_n, y) = 0, \tag{41.5}$$

como también se enuncia y se demuestra el teorema análogo al teorema 1. Para poder enunciarlo, es sólo suficiente que en el enunciado del teorema 1 por x se entienda un punto de un espacio n-dimensional, $x = (x_1, \ldots, x_n) \in R^n$, en particular, $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$.

Teorema 1'. *Supongamos que la función $F(x, y) = F(x_1, \ldots, x_n, y)$ es continua en cierto entorno del punto $(x^{(0)}, y^{(0)})$ y tiene en este entorno la derivada parcial F_y, continua en el punto $(x^{(0)}, y^{(0)})$.*

Si $F(x^{(0)}, y^{(0)}) = 0$ y $F_y(x^{(0)}, y^{(0)}) \neq 0$, existen tales entornos U_x y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$, que para todo $x \in U(x)$ existe una solución, y sólo una,

$$y = f(x) = f(x_1, \ldots, x_n) \in U_y$$

de la ecuación $F(x, y) = 0$[a], con la particularidad de que esta solución es continua en U_x y, además, $y^{(0)} = f(x^{(0)})$.

Si, en adición, en cierto entorno del punto $(x^{(0)}, y^{(0)})$, existen todas las derivadas parciales F_{x_i}, continuas en el punto $(x^{(0)}, y^{(0)})$, entonces en el punto $x^{(0)}$ existen tam-

---

[a] En la fig. 164 se expone el caso en que $n = 2$ y el entorno U_x es rectangular.

bién las derivadas parciales f_{x_i}, $i = 1, 2, \ldots, n$, con la particularidad de que si las derivadas parciales F_{x_i}, $i = 1, 2, \ldots, n$, y F_y son continuas en el entorno del punto $(x^{(0)}, y^{(0)})$, las derivadas parciales f_{x_i} existen y son continuas en cierto entorno del punto $x^{(0)}$.

En este caso las fórmulas para las derivadas parciales de la función implícita, definida por la ecuación (41.5), tienen por expresión

$$\frac{\partial y}{\partial x_i} = -\frac{\dfrac{\partial F}{\partial x_i}}{\dfrac{\partial F}{\partial y}}, \quad i = 1, 2, \ldots, n.$$

Ejercicios. 1. Enúnciense las condiciones, bajo las cuales la función $f(x)$, definida por la ecuación $F(x, y) = 0$ (teorema 1), tiene en el punto (x_0, y_0) derivadas continuas de orden hasta n inclusive. Hállense las fórmulas para $f''(x_0)$ y $f'''(x_0)$.

2. Sirviéndose del teorema 1 y las respuestas de los ejercicios anteriores, hállense las condiciones suficientes para que exista la función $x = \varphi(y)$ que sea inversa respecto de $y = f(x)$ y que tenga en el punto y_0 derivadas continuas de orden hasta n inclusive. Demuéstrese que

$$\frac{d^2x}{dy^2} = -\frac{f''(x)}{[f'(x)]^3}, \quad \frac{d^3x}{dy^3} = \frac{3[f''(x)]^2 - f'(x)f'''(x)}{[f'(x)]^5}.$$

3. Hállense $\dfrac{dy}{dx}$ y $\dfrac{d^2y}{dx^2}$, si y es una función definida por la ecuación $\cos x^2y^2 + xy = 2$.

41.2. PRODUCTOS DE LOS CONJUNTOS

Antes de considerar la cuestión referente a la resolubilidad de los sistemas de ecuaciones, introduzcamos unas nociones adicionales.

Sea R_x^n un espacio euclídeo n-dimensional cuyos puntos se designarán mediante $x = (x_1, \ldots, x_n)$, sea R_y^m un espacio euclídeo m-dimensional cuyos puntos se designarán mediante $y = (y_1, \ldots, y_m)$ y sea R_{xy}^{n+m} un espacio euclídeo $(n + m)$-dimensional de los puntos

$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m).$$

Definición 1. *Supongamos que $A \subset R_x^n$ y $B \subset R_y^m$. El conjunto de tales puntos (x, y) del espacio R_{xy}^{n+m} que $x \in A$ e $y \in B$ se llama producto*[*)] *de los conjuntos A y B y se denota por $A \times B$* (véase el p.1.2*). *De este modo,*

$$A \times B = \{(x, y) : x \in A, y \in B\}.$$

Ejemplos. 1. Si $A = R_x^n$, $B = R_y^m$ entonces

$$A \times B = R_x^n \times R_y^m = R_{xy}^{n+m}.$$

2. Supongamos que $n = 2$ y A es un círculo; $m = 1$ y B es un segmento. En este caso $A \times B$ es un cilindro circular recto (fig. 165).

[*)] Se emplea también el término "*producto cartesiano*".

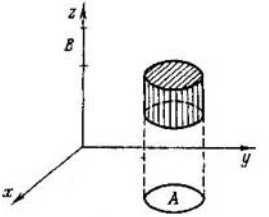


Fig. 165

3. Supongamos que $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R_x^n$ y $A = P(x^{(0)}; \delta_1, \ldots, \delta_n) = \{x: |x_i - x_i^{(0)}| < \delta_i; i = 1, 2, \ldots, n\}$ es un entorno rectangular del punto $x^{(0)}$; supongamos también que $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) \in R_y^m$ y $B = P(y^{(0)}; \eta_1, \ldots, \eta_m) = \{y: |y_j - y_j^{(0)}| < \eta_j, j = 1, 2, \ldots, m\}$ es un entorno rectangular del punto $y^{(0)}$. En este caso

$$A \times B = \{(x, y): |x_i - x_i^{(0)}| < \delta_i, \quad i = 1, 2, \ldots, n;$$
$$|y_j - y_j^{(0)}| < \eta_j, \quad j = 1, 2, \ldots, m\} =$$
$$= P((x^{(0)}, y^{(0)}); \delta_1, \ldots, \delta_n, \quad \eta_1, \ldots, \eta_m) \quad (41.6)$$

es un entorno rectangular del punto $(x^{(0)}, y^{(0)})$.

Lo recíproco es también obvio: puesto que todo entorno rectangular del punto $(x^{(0)}, y^{(0)})$ se escribe mediante la fórmula que está en medio de la igualdad (41.6), siempre puede ser representado como un producto de los entornos rectangulares de los puntos $x^{(0)}$ e $y^{(0)}$.

Ejercicio 4. Demuéstrese que si los conjuntos $A \subset R_x^n$ y $B \subset R_y^m$ son abiertos en los espacios respectivos R_x^n y R_y^m, su producto $A \times B$ será también un conjunto abierto en el espacio R_{xy}^{n+m}.

41.3. FUNCIONES IMPLÍCITAS DEFINIDAS POR UN SISTEMA DE ECUACIONES

Consideraremos las condiciones bajo las cuales el sistema de ecuaciones

$$F_i(x, y) = 0, \quad i = 1, 2, \ldots, m, \quad x \in R^n, \quad y \in R^m, \quad (41.7)$$

o bien, en la forma desarrollada

$$\begin{aligned} &F_1(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \\ &F_2(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &F_m(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \end{aligned} \quad (41.8)$$

es unívocamente resoluble respecto de $y_1, \ldots, y_m$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$ en el que $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$.

Definición 2. *Sea dado un sistema de funciones* $u_i = u_i(t_1, \ldots, t_n)$, $i = 1, 2, \ldots, m$, *que en cierto punto* $t^{(0)}$ *tienen todas las derivadas parciales de primer orden. En este caso una matriz, compuesta de las derivadas parciales de dichas funciones en el punto* $t^{(0)}$

$$\begin{vmatrix} \frac{\partial u_1}{\partial t_1} & \frac{\partial u_1}{\partial t_2} & \cdots & \frac{\partial u_1}{\partial t_n} \\ \frac{\partial u_2}{\partial t_1} & \frac{\partial u_2}{\partial t_2} & \cdots & \frac{\partial u_2}{\partial t_n} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \frac{\partial u_m}{\partial t_1} & \frac{\partial u_m}{\partial t_2} & \cdots & \frac{\partial u_m}{\partial t_n} \end{vmatrix}$$

o, en la forma más breve,

$$\left\| \frac{\partial u_i}{\partial t_j} \right\| \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.$$

se denomina matriz de Jacobi[*)] *del sistema de funciones dado.*

Si $m = n$, el determinante de la matriz de Jacobi se llama *determinante de Jacobi* o *jacobiano* del sistema de funciones $u_1, \ldots, u_n$ según las variables $t_1, \ldots, t_n$ y se designa así[**)]:

$$\frac{\partial(u_1, \ldots, u_n)}{\partial(t_1, \ldots, t_n)}.$$

Veremos en adelante que el jacobiano de un sistema de funciones surge de un modo natural en las más diversas cuestiones de la teoría de la función de varias variables.

Antes de pasar a la exposición del teorema fundamental, mostremos con un ejemplo sencillo (sin profundizarnos en los detalles) la idea de su demostración y señalemos de qué modo surge en sus condiciones el jacobiano del sistema que se considera. Supongamos que en un entorno del punto (x_0, y_0, z_0) se han dado las funciones continuamente derivables F y Φ, con la particularidad de que

$$F(x_0, y_0, z_0) = 0,$$
$$\Phi(x_0, y_0, z_0) = 0.$$

Supongamos también que es preciso resolver el sistema de ecuaciones

$$F(x, y, z) = 0,$$
$$\Phi(x, y, z) = 0$$

---

*) K. Jacobi (1804 — 1851), matemático alemán.

**) Se emplea también la designación $\frac{D(u_1, \ldots, u_n)}{D(t_1, \ldots, t_n)}$.

en cierto entorno del punto indicado, hallando, a partir del sistema, las variables $y = \varphi(x)$ y $z = \psi(x)$ como funciones continuas φ y ψ de la variable x tales que sea $\varphi(x_0) = y_0$, $\psi(x_0) = z_0$. Al resolver con este fin, por ejemplo, la primera ecuación respecto de z, obtendremos $z = f(x, y)$. Al sustituir esta expresión en la segunda y al resolverla respecto de y, tendremos $y = \varphi(x)$. Poniendo $\psi(x) = f[x, \varphi(x)]$, obtendremos la solución buscada

$$y = \varphi(x)$$
$$z = \psi(x).$$

Surge, naturalmente, una pregunta sobre las condiciones que han de cumplirse para que se puedan ejecutar las operaciones indicadas, o, con más precisión, cuándo existen y están unívocamente definidas todas las funciones mencionadas. (Se debe aclarar, por supuesto, ¿dónde, es decir, para qué valores de las variables x e y, están definidas estas funciones? Esta cuestión no será el objeto del análisis detallado inmediato, pues no queremos apartarnos de la idea principal. Volveremos a considerarla en la demostración del teorema 2 de este punto.)

Para que una de las ecuaciones dadas, la primera, por ejemplo, sea resoluble en cierto entorno del punto (x_0, y_0, z_0) respecto de la variable z, es suficiente que (véase el teorema 1′ en el p. 41.1) $\dfrac{\partial F(x_0, y_0, z_0)}{\partial z} \neq 0$. Si $z = f(x, y)$ es la solución correspondiente, entonces, para que la ecuación $\Phi[x, y, f(x, y)] = 0$, que se obtiene como resultado de sustituir dicha solución en la segunda ecuación, sea resoluble respecto de la variable y, resulta suficiente que la derivada parcial total respecto de y en el primer miembro de la igualdad obtenida no se anule en el punto (x_0, y_0), es decir, que en dicho punto se verifique

$$\frac{\partial \Phi}{\partial y} + \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial y} \neq 0.$$

Pero, de acuerdo con el p. 41.1,

$$\frac{\partial f}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}},$$

por consiguiente, al sustituir esta expresión en la desigualdad antecedente, llegamos a que la condición de resolubilidad puede escribirse en la forma

$$\frac{\partial (F, \Phi)}{\partial (y, z)} = \frac{\partial F}{\partial y} \frac{\partial \Phi}{\partial z} - \frac{\partial F}{\partial z} \frac{\partial \Phi}{\partial y} \neq 0 \text{ en el punto } (x_0, y_0, z_0).$$

De esta condición se deduce, obviamente, que en el punto (x_0, y_0, z_0) o bien $\dfrac{\partial F}{\partial z} \neq 0$, o bien $\dfrac{\partial \Phi}{\partial z} \neq 0$, es decir, una de las ecuaciones dadas es resoluble respecto de z.

De este modo, el hecho de que el jacobiano $\dfrac{\partial (F, \Phi)}{\partial (y, z)}$ es distinto de cero en el punto (x_0, y_0, z_0) asegura, para el sistema dado de las ecuaciones, la existencia en cierto

entorno del punto (x_0, y_0, z_0) de una solución en la forma

$$y = \varphi(x)$$
$$z = \psi(x).$$

Enunciemos ahora el teorema fundamental de este párrafo.

Teorema 2. *Supongamos que las funciones* $F_i(x, y) = F_i(x_1, \ldots, x_n, y_1, \ldots, y_m)$, $i = 1, 2, \ldots, m$, *son continuamente derivables en cierto entorno del punto* $(x^{(0)}, y^{(0)})$, *donde* $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)})$. *En este caso, si* $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, *y en el punto* $(x^{(0)}, y^{(0)})$ *el jacobiano* $\dfrac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}$ *no es igual a cero, entonces existen tales entornos* U_x *y* U_y *de los puntos* $x^{(0)}$ *e* $y^{(0)}$ *de los espacios respectivos* R_x^n *y* R_y^m, *que para todo* $x \in U_x$ *existe una única solución*

$$y = f(x) \in U_y$$

del sistema de ecuaciones (41.7):

$$y = f(x) \equiv [y_k = f_k(x_1, \ldots, x_n), \quad k = 1, 2, \ldots, m]^{*)},$$

con la particularidad de que las funciones $f_k(x)$, $k = 1, 2, \ldots, m$, *que forman dicha solución son continuamente derivables en* U_x *y, además,* $f(x^{(0)}) = y^{(0)}$.

Así pues, si se cumplen las suposiciones del teorema, la condición

$$F_i(x, y) = 0, \quad i = 1, 2, \ldots, m, \quad (x, y) \in U_x \times U_y$$

es equivalente a la siguiente

$$y = f(x), \quad x \in U_x, \quad y \in U_y.$$

DEMOSTRACIÓN. Indiquemos ante todo que la afirmación: la solución $y = f(x)$ del sistema de ecuaciones (41.7) satisface la condición $f(x^{(0)}) = y^{(0)}$ proviene, evidentemente y de inmediato, de la afirmación sobre la unicidad de la solución $y = f(x) \in U_y$ para $x \in U_x$ y de las condiciones $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, $x^{(0)} \in U_x$, $y^{(0)} \in U_y$.

Para demostrar el teorema apliquemos el método de la inducción matemática. Para el caso de una ecuación, es decir, cuando $m = 1$, el teorema se ha enunciado en el p. 41.1. Supongamos ahora que es también lícito para $m - 1$ ecuaciones ($m > 1$). Demostremos que en tal caso tendrá lugar también para m ecuaciones.

Mostremos primero que cada una de las ecuaciones (41.8), la última, por ejemplo,

$$F_m(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0,$$

---

*) El sistema de funciones $f_k(x_1, \ldots, x_n)$, $k = 1, 2, \ldots, m$, viene designado con un símbolo $f(x)$, puesto que define una correspondencia determinada: a los puntos de cierto conjunto del espacio R_x^n el sistema indicado les pone en correspondencia los puntos determinados del espacio R_y^m, o, como suele decirse, aplica el conjunto mencionado del espacio R_x^n en el espacio R_y^m.

puede ser resuelta en el entorno del punto $(x^{(0)}, y^{(0)})$ por lo menos respecto de una variable. En efecto, por hipótesis del teorema, en el punto $(x^{(0)}, y^{(0)})$

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)} = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \cdots\cdots & \cdots & \cdots\cdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{vmatrix} \neq 0,$$

y, por ende, en dicho punto por lo menos un elemento en la última fila del jacobiano es distinto de cero. Supongamos, para concretar, que este elemento será el último:

$$\frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_m} \neq 0.$$

De aquí, en virtud del teorema 1′ del p. 41.1, se desprende que la ecuación $F_m(x, y) = 0$ puede ser resuelta respecto de y_m en cierto entorno del punto (x_0, y_0). Daremos una definición más exacta de esta idea. Designemos mediante U el entorno del punto $(x^{(0)}, y^{(0)})$, en el que las funciones F_i, $i = 1, 2, \ldots, m$, son continuamente derivables y pongamos $\tilde{y} \overset{\text{def}}{=} (y_1, \ldots, y_{m-1})$. Entonces, existen un entorno rectangular U^{m+n-1} del punto

$$(x^{(0)}, \tilde{y}^{(0)}) = (x_1^0, \ldots, x_n^{(0)}, y_1^{(0)}, \ldots, y_{m-1}^{(0)}) \tag{41.9}$$

y un entorno U^1 del punto $y_m^{(0)}$ tales que $U^{m+n-1} \times U^1 \subset U$, y se tiene la única función

$$y_m = \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}), \tag{41.10}$$

que está definida en U^{m+n-1} y que satisface las siguientes condiciones: si

$$(x, \tilde{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) \in U^{m+n-1},$$

entonces

$$\varphi(x, \tilde{y}) = \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) \in U^1, \tag{41.11}$$

$$F_m(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}, \varphi(x, \tilde{y})) = 0. \tag{41.12}$$

Además, de acuerdo con el mismo teorema 1′, la función $\varphi(x, \tilde{y})$ es continuamente derivable en U^{m+n-1} y

$$\varphi(x^{(0)}, \tilde{y}^{(0)}) = y_m^{(0)}. \tag{41.13}$$

En este caso, si $(x, \tilde{y}) \in U^{m+n-1}$ e $y_m \in U^1$, entonces el sistema (41.8) es equivalente al sistema

$$F_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1, \quad y_m = \varphi(x, y). \tag{41.14}$$

Sustituyamos en las primeras $m - 1$ ecuaciones del sistema (41.14) la expresión (41.10). Al introducir las designaciones

$$\Phi_i(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = F_i(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}, \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}))$$
$$i = 1, 2, \ldots, m-1. \quad (41.15)$$

obtendremos, en este caso, el siguiente sistema de $m-1$ ecuaciones con $m+n-1$ incógnitas:

$$\begin{aligned} &\Phi_1(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\Phi_{m-1}(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = 0. \end{aligned} \quad (41.16)$$

Además, para $(x, \bar{y}) \in U^{m+n-1}$, $y_m \in U^1$ el sistema de ecuaciones

$$\begin{aligned} &\Phi_j(x, \bar{y}) = 0, \quad j = 1, 2, \ldots, m-1, \\ &y_m = \varphi(x, \bar{y}) \end{aligned} \quad (41.17)$$

es equivalente al sistema (41.14).

Mostremos que el sistema (41.16) satisface las condiciones que se diferencian de las que son satisfechas por el sistema (41.8), sólo en que $m-1$ se ha sustituido por m. En efecto, las funciones Φ_k, $k = 1, 2, \ldots, m-1$, son continuamente derivables en el entorno U^{m+n-1} como composición de las funciones continuamente derivables. De las condiciones $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, y (41.15), (41.13) proviene que $\Phi_k(x^{(0)}, y^{(0)}) = 0$, $k = 1, 2, \ldots, m-1$.

Demostremos que en el punto $(x^{(0)}, \bar{y}^{(0)})$ (véase (41.9))

$$\frac{\partial(\Phi_1, \ldots, \Phi_{m-1})}{\partial(y_1, \ldots, y_{m-1})} \neq 0.$$

Indiquemos previamente que de (41.10) y (41.15) se deduce que

$$\frac{\partial \Phi_i}{\partial y_k} = \frac{\partial F_i}{\partial y_k} + \frac{\partial F_i}{\partial y_m} \frac{\partial \varphi}{\partial y_k}, \quad i, k = 1, 2, \ldots, m-1, \quad (41.18)$$

y de (41.12), que

$$\frac{\partial F_m}{\partial y_k} + \frac{\partial F_m}{\partial y_m} \frac{\partial \varphi}{\partial y_k} = 0, \quad k = 1, 2, \ldots, m-1. \quad (41.19)$$

Ahora, en el determinante $\dfrac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}$ agreguemos a la k-ésima columna la última columna multiplicada por $\dfrac{\partial \varphi}{\partial y_k}$, $k = 1, \ldots, m-1$, de lo cual, como se sabe, el valor del determinante no cambiará. Por eso, utilizando (41.18) y (41.19) y desarrollando el determinante obtenido por los elementos de la última fila, obtenemos

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}_{(x^{(0)}, y^{(0)})} =$$

$$= \begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} + \dfrac{\partial F_1}{\partial y_m}\dfrac{\partial \varphi}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_{m-1}} + \dfrac{\partial F_1}{\partial y_m}\dfrac{\partial \varphi}{\partial y_{m-1}} & \dfrac{\partial F_1}{\partial y_m} \\ \cdots\cdots & & \cdots\cdots & \cdots\cdots \\ \dfrac{\partial F_m}{\partial y_1} + \dfrac{\partial F_m}{\partial y_m}\dfrac{\partial \varphi}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_{m-1}} + \dfrac{\partial F_m}{\partial y_m}\dfrac{\partial \varphi}{\partial y_{m-1}} & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix}_{(x^{(0)}, y^{(0)})} =$$

$$= \begin{vmatrix} \dfrac{\partial \Phi_1}{\partial y_1} & \cdots & \dfrac{\partial \Phi_1}{\partial y_{m-1}} & \dfrac{\partial F_1}{\partial y_m} \\ \cdots\cdots & & \cdots\cdots & \cdots\cdots \\ \dfrac{\partial \Phi_{m-1}}{\partial y_1} & \cdots & \dfrac{\partial \Phi_{m-1}}{\partial y_{m-1}} & \dfrac{\partial F_{m-1}}{\partial y_m} \\ 0 & \cdots & 0 & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix}_{(x^{(0)}, \tilde{y}^{(0)})} =$$

$$= \frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_m} \left| \frac{\partial (\Phi_1, \ldots, \Phi_{m-1})}{\partial (y_1, \ldots, y_{m-1})} \right|_{(x^{(0)}, \tilde{y}^{(0)})}$$

y como el primer miembro de la igualdad es distinto de cero, será no nulo también el segundo miembro, de donde

$$\left| \frac{\partial (\Phi_1, \ldots, \Phi_{m-1})}{\partial (y_1, \ldots, y_{m-1})} \right|_{(x^0, \tilde{y}^0)} \neq 0.$$

En vista de que las condiciones, análogas a las condiciones para las funciones F_i, $i = 1, 2, \ldots, m$, se cumplen para las funciones Φ_i, $i = 1, 2, \ldots, m-1$, y, de acuerdo con la suposición sobre la inducción, el sistema de ecuaciones (41.16) es unívocamente resoluble respecto de las variables $y_1, \ldots, y_{m-1}$ en cierto entorno del punto $(x^{(0)}, \tilde{y}^{(0)})$. Con más precisión, sea U^{m+n-1} un entorno rectangular del punto $(x^{(0)}, \tilde{y}^{(0)})$ obtenido como resultado de la resolución de la ecuación $F_m = 0$ respecto de la variable y_m. Desarrollémoslo en un producto de los entornos rectangulares U'_x y $U'_{\tilde{y}}$ de los puntos $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ e $\tilde{y}^{(0)} = (y_1^{(0)}, \ldots, y_{m-1}^{(0)})$ en los espacios respectivos R_x^n y R_y^{m-1} (aquí $y = (y_1, \ldots, y_{m-1})$): $U^{m+n-1} = U'_x \times U'_{\tilde{y}}$. Entonces, existen un entorno $U_x \subset U'_x$ del punto $x^{(0)}$, un entorno $U_{\tilde{y}} \subset U'_{\tilde{y}}$ del punto $\tilde{y}^{(0)}$ y el único sistema de funciones

$$\begin{aligned} y_1 &= f_1(x) = f_1(x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{m-1} &= f_{m-1}(x) = f_{m-1}(x_1, \ldots, x_n), \end{aligned} \tag{41.20}$$

que están definidas en el conjunto U_x y que satisfacen las siguientes condiciones: si $x \in U_x$, se tiene

$$(f_1(x), \ldots, f_{m-1}(x)) \in U_{\bar{y}}. \tag{41.21}$$

y en U_x las funciones (41.20) son continuamente derivables y satisfacen el sistema de ecuaciones (41.16):

$$\Phi_i(x_1, \ldots, x_n, f_1(x), \ldots, f_{m-1}(x)) = 0, \qquad i = 1, 2, \ldots, m-1. \tag{41.22}$$

Es importante notar que en virtud de la unicidad de la solución (41.20) del sistema (41.16) para $x \in U_x$, $\bar{y} \in U_{\bar{y}}$ e $y_m \in U^1$, el sistema de ecuaciones

$$\begin{aligned} y_k &= f_k(x), \quad k = 1, 2, \ldots, m-1, \\ y_m &= \varphi(x, y) \end{aligned} \tag{41.23}$$

es equivalente al sistema (41.17).

Sustituyendo las expresiones (41.20) en (41.10), obtendremos una función de x, definida en U_x; designémosla con f_m:

$$y_m = \varphi(x_1, \ldots, x_n, f_1(x), \ldots, f_{m-1}(x)) = f_m(x_1, \ldots, x_n) = f_m(x). \tag{41.24}$$

Señalemos que el sistema de funciones

$$y_k = f_k(x_1, \ldots, x_n), \quad k = 1, 2, \ldots, m, \tag{41.25}$$

(véanse (41.20) y (41.24)) es precisamente el sistema buscado de funciones que satisface las exigencias enunciadas en el teorema. En efecto, sea $U_y = U_{\bar{y}} \times U^1$; entonces si $x \in U_x$, de conformidad con (41.21) y (41.11), $f(x) = (f_1(x), \ldots, f_m(x)) \in U_y$. De (41.15), (41.22), (41.24) y (41.12) se deduce que $F_i(x, f(x)) = 0$, $i = 1, 2, \ldots, m$, para cualquier $x \in U_x$. Debido al teorema 1' y la suposición de inducción, las funciones (41.10) y (41.20) y, por lo tanto, la función (41.24), son continuamente derivables.

Hemos demostrado de este modo que la aplicación $f(x)$, definida por las funciones (41.25), es una solución continuamente derivable del sistema de ecuaciones (41.8) en el conjunto U_x, y si, en este caso, $x \in U_x$, entonces $y = f(x) \in U_y$. Observemos, además, que si $x \in U_x$, el sistema (41.25) es equivalente al sistema (41.23).

Resta demostrar la unicidad de la solución del sistema de ecuaciones (41.8). Con este fin representemos en forma del esquema que sigue los pasos (realizados en el transcurso de la demostración) de unos sistemas de ecuaciones a los otros, equivalentes a los primeros, es decir, a aquellos que tienen exactamente las mismas soluciones:

$$\begin{aligned} &F_i(x, y) = 0, \quad i = 1, 2, \ldots, m. \\ &\Updownarrow \\ &F_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1, \\ &y_m = \varphi(x, y). \\ &\Updownarrow \\ &\Phi_j(x, \bar{y}) = F_j(x, \bar{y}, \varphi(x, \bar{y})), \quad j = 1, 2, \ldots, m-1, \\ &\Phi_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1, \\ &y_m = \varphi(x, y). \\ &\Updownarrow \\ &y_j = f_j(x), \quad j = 1, 2, \ldots, m-1, \end{aligned}$$

$$y_m = \varphi(x, \tilde{y}). \qquad f_m(x) \equiv \varphi(x, f_1(x), \ldots, f_{m-1}(x))$$
$$\Updownarrow$$
$$y_i = f_i(x), \quad i = 1, 2, \ldots, m.$$

Las flechas dobles significan la equivalencia de los sistemas de ecuaciones en consideración, la cual en todo caso tiene lugar para $x \in U_x$, $y \in U_y$. De esta equivalencia deriva precisamente la unicidad de la solución (41.25) del sistema (41.8) en los entornos que se consideran, de lo cual, según se ha observado anteriormente, en virtud de la condición $F_i(x^{(0)}, y^{(0)}) = 0, i = 1, 2, \ldots, m$, se deduce que $f(x^{(0)}) = y^{(0)}$. □

El teorema demostrado sobre las funciones implícitas es uno de los más importantes teoremas del análisis matemático y tiene toda una serie de aplicaciones en diferentes apartados de éste. Con algunas de ellas nos encontraremos en los capítulos ulteriores de nuestro curso. El teorema citado es un "teorema puro de existencia": tanto su enunciado, como la demostración aducida no dan origen, en el caso general, a un método concreto de resolución del sistema (41.8). Por ejemplo, si todas las $F_k, k = 1, 2, \ldots, m$, en el sistema citado de ecuaciones son funciones elementales, entonces, siguiendo el esquema de la demostración del teorema, no tendremos éxito en "hallar (se trata de un caso general) en la forma explícita" todas aquellas funciones cuya existencia se ha utilizado en la demostración mencionada, y obtener una solución del sistema que tenga también la forma de las funciones elementales. Realmente, en este caso la solución del sistema de ecuaciones (41.8), la que existe en virtud del teorema citado, no es, en el caso general, una lista de las funciones elementales (si incluso dicho sistema consta de una sola ecuación).

Naturalmente, si las funciones F_k son elementales y, por consiguiente, vienen definidas mediante ciertas fórmulas, la solución del sistema (41.8) puede hallarse con cualquier grado de exactitud, es decir, en principio, las tablas de los valores de estas soluciones pueden ser compuestas con cualquier grado de exactitud. En realidad, sin embargo, la exactitud, con la que se calculan las soluciones, se determina, desde luego, por el objetivo concreto, en aras del cual se resuelve el sistema en consideración. El propio teorema 2 en este caso nos da una certeza objetiva de que, al realizar correctamente los cálculos correspondientes, calculamos, de hecho, la solución buscada del sistema. No nos detendremos en los métodos numéricos que se emplean para solucionar los sistemas de ecuaciones; sólo algunas de las cuestiones de la resolucion numerica de las ecuaciones se consideran en "Complemento" al final de este tomo.

Indiquemos como una circunstancia esencial que el teorema 2, al igual que todos los teoremas de este tipo, proporciona métodos cualitativos, en el caso dado, para estudiar las propiedades de las soluciones del sistema de ecuaciones.

Resulta interesante notar que las derivadas parciales de la solución del sistema (41.8) se expresan con facilidad, si se cumplen las condiciones del teorema 2, en forma explícita a través de las derivadas parciales de las funciones F_k, $k = 1, 2, \ldots, m$. Efectivamente, con el fin de hallar la derivada parcial $\dfrac{\partial y_j}{\partial x_i}$, se deben derivar las igualdades (41.8) respecto de x_i, considerándolas identidades según $x_1, \ldots, x_n$, es decir, sustituyendo en ellas sus soluciones $y_j = y_j(x_1, \ldots, x_n), j = 1, \ldots, m$. En este caso obtendremos

$$\frac{\partial F_k}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j} \frac{\partial y_j}{\partial x_i} = 0, \quad k = 1, 2, \ldots, m.$$

Este sistema de ecuaciones, lineales respecto de $\frac{\partial y_j}{\partial x_i}$, en virtud de que su determinante en el punto considerado es distinto de cero:

$$\frac{\partial (F_1, \ldots, F_m)}{\partial (y, \ldots, y_m)} \neq 0,$$

tiene una solución, y sólo una, la cual puede hallarse, por ejemplo, según la regla de Cramer*).

Si es preciso hallar todas las derivadas

$$\frac{\partial y_j}{\partial x_i}, \quad i = 1, 2, \ldots, n \quad j = 1, 2, \ldots, m,$$

conviene calcular las diferenciales de ambos miembros de las identidades mencionadas arriba (41.8). Haciendo uso de la invariación de la forma de la primera diferencial respecto de la elección de las variables, obtendremos

$$\sum_{j=1}^{n} \frac{\partial F_k}{\partial x_i} dx_i + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j} dy_j = 0, \quad k = 1, 2, \ldots, m.$$

En vista de la misma condición $\frac{\partial (F_1, \ldots, F_m)}{\partial (y_1, \ldots, y_m}$ $\neq 0$, este sistema de las ecuaciones, lineales respecto de $dy_1, \ldots, dy_m$, tiene una solución y ésta es única. Si la hallamos, el coeficiente de dx_i en la expresión para dy_j será precisamente la derivada parcial $\frac{\partial y_j}{\partial x_i}$.

Los dos métodos son aplicables también para el cálculo de las derivadas de órdenes superiores de las funciones $y_j(x_1, \ldots, x_n)$ que son soluciones del sistema de las ecuaciones (41.8) (por ejemplo, bajo el supuesto de que todas las funciones F_k, $k = 1, 2, \ldots, m$, tienen derivadas continuas de órdenes correspondientes). Aplicando el método de diferenciales, se debe recordar, por supuesto, que las diferenciales de orden superior al primero, si se expresan en términos de las diferenciales de las funciones, son más complejas, en lo que se refiere a su expresión, en comparación con el caso cuando ellas se expresan sólo en términos de las diferenciales de las variables independientes (véase el p. 21.2).

Las derivadas de órdenes superiores de las funciones $y_j(x_1, \ldots, x_n)$ pueden obtenerse también, por derivación sucesiva, a partir de las expresiones para las prime-

*) G. Cramer (1704 — 1752), matemático suizo.

ras derivadas $\frac{\partial y_j}{\partial x_i}$, determinadas según las fórmulas de Cramer a base del sistema de ecuaciones citado anteriormente

$$\frac{\partial F_k}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j} \frac{\partial y_j}{\partial x_i} = 0, \quad k = 1, 2, \ldots, m,$$

en forma de una razón entre dos determinantes. Dicha razón puede derivarse tantas veces cuantas veces son derivables las funciones F_k, $k = 1, \ldots, m$. En este caso, si todas las derivadas de las funciones F_k, $k = 1, \ldots, m$, de orden hasta r inclusive son continuas, lo serán también todas las derivadas parciales de las funciones $y_j(x_1, \ldots, x_n)$, $j = 1, \ldots, m$, de orden hasta el mismo r.

Un conjunto (llamado también a menudo una clase) de todas las funciones que son r veces continuamente derivables en la región G se denota con $C^r(G)$. De este modo: *si, adicionalmente a las condiciones del teorema 2, $F_k \in C^r(U)$, $k = 1, \ldots, m$, donde U es un cierto entorno del punto $(x^{(0)}, y^{(0)})$, las soluciones $y_j = y_j(x_1, \ldots, x_n)$ del sistema de ecuaciones* (41.7) *pertenece también a la clase $C^r(U_x)$ en cierto entorno U_x del punto $x^{(0)}$.*

Ejercicios. 5. ¿Bajo qué condiciones impuestas sobre f y g, la ecuación $y = xf(z) + g(z)$ define, en cierto entorno U del punto (x_0, y_0), la función $z(x, y) \in C^2(U)$? Demuéstrese que, cumplidas estas condiciones, para cualesquiera $(x, y) \in U$

$$z_y^2 z_{xx} - 2z_x z_y z_{xy} + z_x^2 z_{yy} = 0.$$

6. Sea dado un sistema de ecuaciones

$$uf'(v) = [y - f(v)]^2$$
$$(x + v)f'(v) = y - f(v).$$

Hállense las condiciones, impuestas sobre la función f, para las cuales este sistema define en cierto entorno U del punto (x_0, y_0) las funciones $u = u(x, y)$, $v = v(x, y)$ de la clase $C^1(U)$. Demuéstrese que en este caso $u_x u_y = u$ en todo punto de U.

41.4. APLICACIONES

En este punto se estudiarán las aplicaciones $f: X \to R^m$, $X \subset R^n$, es decir, aplicaciones de tal índole que a todo punto $x = (x_1, \ldots, x_n)$ del conjunto X, dispuesto en el espacio puntual aritmético n-dimensional R^n (véase el p. 18.1), le ponen en correspondencia el punto $y = (y_1, \ldots, y_m)$ del espacio puntual aritmético m-dimensional R^m. De este modo, $f: (x_1, \ldots, x_n) \to (y_1, \ldots, y_m)$, $(x_1, \ldots, x_n) \in X$. Obviamente, la definición de tal aplicación f es equivalente a la definición de m funciones $f_j: X \to R$ tales que sea $f_j: x \mapsto y_j$, $j = 1, \ldots, m$, $x \in X$, $y_j \in R$. Estas funciones

$$f_j(x) = f_j(x_1, \ldots, x_n), \quad i = 1, 2, \ldots, m, \quad x \in X, \tag{41.26}$$

se denominan *funciones coordenadas de la aplicación f* y se escribe

$$f = (f_1, \ldots, f_m).$$

A las aplicaciones que se consideran se generaliza el concepto de continuidad.

Definición 3. *Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier entorno $V(y^{(0)})$ del punto $y^{(0)} = f(x^{(0)})$ existe tal entorno $U(x^{(0)})$ del punto $x^{(0)}$ que*

$$f(U(x^{(0)}) \cap X \subset V(y^{(0)}).$$

Por cuanto en todo entorno de un punto*) se contiene su entorno esférico, dicha definición es equivalente a la siguiente.

Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier ε-entorno del punto $y^{(0)} = f(x^{(0)})$ existe tal δ-entorno del punto $x^{(0)}$ que

$$f(U(x^{(0)}, \delta) \cap X) \subset U(y^{(0)}, \varepsilon).$$

Esto, a su vez, puede ser parafraseado, con ayuda de las desigualdades, de la manera siguiente.

Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier $\varepsilon > 0$ existe tal $\delta > 0$ que para todos los puntos $x \in X$, que satisfacen la condición $\rho(x, x^{(0)}) < \delta$, se cumple la desigualdad

$$\rho(f(x), f(x^{(0)})) < \varepsilon.$$

La definición de continuidad puede enunciarse también en términos de las sucesiones.

Definición 3'. *Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$ si para cualquier sucesión $x^{(k)} \in X$, $k = 1, 2, \ldots$, tal que $\lim\limits_{k \to \infty} x^{(k)} = x^{(0)}$, tiene lugar*

$$\lim_{k \to \infty} f(x^{(k)}) = f(x^{(0)}).$$

La equivalencia de estas dos definiciones se demuestra por analogía con la demostración de la equivalencia de las definiciones del límite de funciones según Cauchy y según Heine. Demos a conocer esta demostración.

Supongamos que la aplicación *f es continua en el punto $x^{(0)}$* en el sentido de la definición 3, $x^{(k)} \in X$, $k = 1, 2, \ldots$, y

$$\lim_{k \to \infty} x^{(k)} = x^{(0)}. \tag{41.27}$$

Prefijemos $\varepsilon > 0$. Para ε existe un $\delta > 0$ tal que si $x \in X$, $\rho(x, x^{(0)}) < \delta$, se verifica la desigualdad $\rho(f(x), f(x^{(0)})) < \varepsilon$.

En virtud de la condición (41.27), existe tal número k_0 que para cualesquiera $k \geqslant k_0$ tenemos $x^{(k)} \in U(x^{(0)}, \delta)$ y, por consiguiente, $\rho(f(x^{(k)}), f(x^{(0)})) < \varepsilon$. Esto precisamente significa que $\lim\limits_{k \to \infty} f(x^{(k)}) = f(x^{(0)})$.

Supongamos ahora que la aplicación f es continua en el punto $x^{(0)}$ en el sentido de la definición 3' y que las condiciones de la definición 3 no están cumplidas, es decir, existe tal $\varepsilon_0 > 0$, que para cualquier $\delta > 0$ existe un $x_\delta \in U(x^{(0)}, \delta) \cap X$, para el

---

*) Recordemos que se llama entorno de un punto cualquier conjunto abierto que contiene dicho punto (véase la definición 14 en el p. 18.2).

cual $\rho(f(x_k), f(x^{(0)})) \geqslant \varepsilon_0$. Tomando sucesivamente $\delta = \frac{1}{k}$, $k = 1, 2, \ldots$, y poniendo, para abreviar, $x^{(k)} = x_{1/k}$, obtenemos $x^{(k)} \in U\left(x^{(0)}, \frac{1}{k}\right) \cap X$, es decir, $\rho(x^{(k)}, x^{(0)}) < \frac{1}{k}$. Por consiguiente, $\lim_{k\to\infty} x^{(k)} = x^{(0)}$ y $x^{(k)} \in X$; sin embargo, $\rho(f(x^{(k)}), f(x^{(0)})) \geqslant \varepsilon_0$ y, de este modo, la sucesión $\{f(x^{(k)})\}$ no tiene el punto $f(x^{(0)})$ en calidad de su límite. La contradicción obtenida demuestra la afirmación enunciada. □

Lema 1. *La aplicación $f = (f_1, \ldots, f_m)$: $X \to R^m$, $X \subset R^n$, es continua en el punto $x^{(0)}$ cuando, y sólo cuando, en dicho punto son continuas todas las funciones coordenadas $f_1, \ldots, f_m$.*

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que la aplicación f es continua en el punto $x^{(0)} \in X$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) \stackrel{\text{def}}{=} f(x^{(0)})$. De acuerdo con la definición 3, para todo entorno $V(y^{(0)})$ del punto $y^{(0)}$, en particular, para cada uno de sus entornos cúbicos (véase el p. 18.1)

$$P(y^{(0)}, \varepsilon) = \{y: |y_l - y_l^{(0)}| < \varepsilon\}$$

existe tal entorno $U(x^{(0)})$ del punto $x^{(0)}$ que

$$f(U(x^{(0)}) \cap X) \subset P(y^{(0)}, \varepsilon).$$

Por consiguiente, para todo $x \in U(x^{(0)}) \cap X$ se verifican las desigualdades

$$|f_j(x) - y_j^{(0)}| < \varepsilon, \quad j = 1, 2, \ldots, m.$$

Esto es testimonio de que todas las funciones coordenadas $f_1, \ldots, f_m$ son continuas en el punto $x^{(0)}$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Supongamos que todas las funciones coordenadas $f_1, \ldots, f_m$ son continuas en el punto $x^{(0)} \in X$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) = f(x^{(0)})$ y está dado un entorno $V(y^{(0)})$ del punto $y^{(0)}$. En este caso existe tal $\varepsilon > 0$ que el entorno ε-cúbico $P(y^{(0)}, \varepsilon)$ del punto $y^{(0)}$ está contenido en $V(y^{(0)})$,

$$P(y^{(0)}, \varepsilon) \subset V(y^{(0)}).$$

Por ser continua cada una de las funciones f_j, $j = 1, 2, \ldots, m$, en el punto $x^{(0)}$ existen tales entornos $U_j = U(x^{(0)})$ que para $x \in U_j \cap X$ se verifica la desigualdad

$$|f_j(x) - y_j^{(0)}| < \varepsilon. \tag{41.28}$$

Pongamos $U = \bigcap_{j=1}^{m} U_j$. Entonces, U, siendo una intersección de un número finito de los conjuntos abiertos U_j, será conjunto abierto, con la particularidad de que como todos los U_j contenían el punto $x^{(0)}$, lo contiene también U. De este modo, el conjunto U es un entorno del punto $x^{(0)}$. Además, si $x \in U \cap X$, entonces para todo $j = 1, 2, \ldots, m$, se verifican las desigualdades (41.28). Esto quiere decir que

$$f(x) \in P(y^{(0)}, \varepsilon),$$

y, por lo tanto, $f(x) \in V(y^{(0)})$. Así pues, para un entorno arbitrario $V(y^{(0)})$ se ha encontrado tal entorno U del punto $x^{(0)}$ que

$$f(U \cap X) \subset V(y^{(0)}). \square$$

El lema 1 muestra, en particular, que las definiciones de las aplicaciones continuas de un segmento introducidas al considerar el concepto de curva en el p. 16.1 (cuando el segmento se aplica en un espacio tridimensional) y en el p. 18.2 (cuando el segmento se aplica en un espacio euclídeo arbitrario n-dimensional) como aplicaciones cuyas funciones coordenadas son continuas, son equivalentes a la definición de las aplicaciones continuas de un segmento como aplicaciones de tal género que en todo punto del segmento satisfacen las condiciones de la definición 3 de este párrafo.

La aplicación: $f: X \to R_y^m$ $X \subset R_x^n$ se llama *continua en el conjunto* X, si es continua en todo punto del conjunto X.

Lema 2. *La aplicación f de un conjunto abierto del espacio R_x^n en el espacio R_y^m es continua en dicho conjunto cuando, y sólo cuando, la preimagen de cada conjunto abierto del espacio R_y^m, realizándose la aplicación f, será un conjunto abierto del espacio R_x^n.*

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que f aplica continuamente el conjunto abierto $G \subset R_x^n$ en el espacio R_y^m y sea U un conjunto abierto del espacio R_y^m: $U \subset R_y^m$. Mostremos que la preimagen $f^{-1}(U)$ de este conjunto es un conjunto abierto en el espacio R_x^n. Si el conjunto $f^{-1}(U)$ es vacío, la afirmación queda obvia, puesto que todo conjunto vacío es abierto.

Supongamos que el conjunto $f^{-1}(U)$ no es vacío, es decir, existe un punto $x^{(0)} \in f^{-1}(U)$ y, por lo tanto, $f(x^{(0)}) \in U$. Como U es un conjunto abierto, será un entorno del punto $y^{(0)} = f(x^{(0)})$. Por ello, por ser la aplicación f continua en el punto $x^{(0)}$ (véase la definición 3°), existe tal entorno U_x de este punto que $f(U_x \cap G) \subset U$, por consiguiente, $U_x \cap G \subset f^{-1}(U)$. Como el conjunto $U_x \cap G$, representando una intersección de dos conjuntos abiertos U_x y G, es abierto y puesto que $x^{(0)} \in U_x \cap G$, entonces $x^{(0)}$ es un punto interior del conjunto $f^{-1}(U)$.

De este modo, todo punto de la preimagen del conjunto abierto U es un punto interior de esta preimagen, quiere decir, la preimagen es un conjunto abierto.

DEMOSTRACIÓN DE LA SUFICIENCIA. Sea f una aplicación del conjunto abierto G del espacio R_x^n en R_y^m y supongamos que, al realizarse esta aplicación, la preimagen de cada conjunto abierto en el espacio R_y^m es un conjunto abierto en R_x^n. Sea $x^{(0)} \in G$. Mostremos que la aplicación f es continua en el punto $x^{(0)}$.

Sea U_y cierto entorno del punto $y^{(0)} = f(x^{(0)})$. Como la preimagen $f^{-1}(U_y)$ del conjunto abierto U_y es, por hipótesis, un conjunto abierto, y, evidentemente, $x^{(0)} \in f^{-1}(U_y) \subset G$, entonces el conjunto $U_x = f^{-1}(U_y)$ es un entorno del punto $x^{(0)}$ y, además, $f(U_x) = U_y$. De aquí se deduce de inmediato la continuidad de la aplicación f en el punto $x^{(0)}$ (véase la definición 3). $\square$

Ejemplo. Examinemos la aplicación $f: R^2 \to R$, definida por la fórmula $f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$. De acuerdo con el lema 2, la preimagen del conjunto abierto $(-\infty, 0)$, es decir, un conjunto de puntos (x, y) que satisfacen la desigualdad $\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1$ (y, por consiguiente, forman el interior de una elipse), como también la preimagen del conjunto abierto $(0, +\infty)$, o sea, un conjunto de tales puntos (x, y) que $\frac{x^2}{a^2} + \frac{y^2}{b^2} > 1$ (estos puntos forman el exterior de una elipse), son conjuntos abiertos.

En general, si $f: R^n \to R$ es una función continua en R^n, para cualquier número $a \in R$ los conjuntos $\{x: f(x) < a,\ x \in R^n\}$ y $\{x: f(x) > a,\ x \in R^n\}$ son conjuntos abiertos, siendo preimágenes de los conjuntos abiertos $(-\infty, a)$ y $(a, +\infty)$.

El teorema de Weierstrass en el cual se demuestra que las funciones continuas en un compacto son acotadas y pueden alcanzar sus cotas inferior y exterior se extiende también al caso de las aplicaciones continuas. Con más precisión, resulta válida la siguiente afirmación.

Lema 3. *Sea $f: A \to R^m$, $A \subset R^n$ una aplicación continua de un compacto A en el espacio R^m. El conjunto $f(A)$ es también un compacto en este caso.*

En la forma más breve: *la preimagen continua de un compacto es un compacto.*

DEMOSTRACIÓN. Sea $y^{(k)} \in f(A)$ una sucesión arbitraria de los puntos pertenecientes a $f(A)$. Por definición de la preimagen de un conjunto para la aplicación dada, existe tal punto $x^{(k)} \in A$, que $f(x^{(k)}) = y^{(k)}$, cualquiera que sea $k = 1, 2, \ldots$. Como A es un compacto, de la sucesión $\{x^{(k)}\}$ se puede separar una subsucesión convergente $\{x^{(k_s)}\}$ cuyo límite $x^{(0)}$ pertenece al compacto A: $\lim\limits_{s \to \infty} x^{(k_s)} = x^{(0)} \in A$.

Por ser la función f continua en el punto $x^{(0)}$, tenemos

$$\lim_{s \to \infty} f(x^{(k_s)}) = f(x^{(0)}), \text{ es decir, } \lim_{s \to \infty} y^{(k_s)} = f(x^{(0)}) \in f(A).$$

De este modo, de toda sucesión de puntos, pertenecientes al conjunto $f(A)$, se puede separar una sucesión convergente cuyo límite pertenece a dicho conjunto. Esto precisamente significa que $f(A)$ es un compacto. □

OBSERVACIÓN. Del lema 3 se infiere el teorema demostrado anteriormente de que una función real continua en un compacto alcanza sus cotas inferior y exterior (véase el p. 19.5). En efecto, de conformidad con el lema 3, el conjunto de valores de tal función es un compacto en una recta numérica, mientras que todo compacto tiene en una recta numérica los puntos finitos maximal y minimal. Esto proviene de que un compacto es un conjunto acotado y, por ende, tiene la cota superior (inferior) finita la cual, en virtud de su definición, es un punto adherente del conjunto. Ya que el compacto está cerrado, el punto pertenece a él y es, evidentemente, su punto maximal (minimal).

La noción de continuidad uniforme se generaliza también para el caso de las aplicaciones.

Definición 4. *La aplicación f del conjunto $X \subset R_x^n$ en el espacio R_y^m se llama uniformemente continua, si para cualquier $\varepsilon > 0$ existe tal $\delta = \delta(\varepsilon) > 0$, que para cualesquiera puntos $x' \in X$ y $x'' \in X$, que satisfacen la condición $\rho(x', x'') < \delta$, se verifica la desigualdad $\rho(f(x'), f(x'')) < \varepsilon$.*

Para las aplicaciones es lícita también una afirmación análoga al teorema de Cantor (véase el p. 19.6) para las funciones continuas.

Lema 4. *Una aplicación continua de un compacto es uniformemente continua.*

DEMOSTRACIÓN. Hagamos uso del mismo método que se aplicó al demostrar el teorema de Cantor sobre la continuidad uniforme de las funciones reales continuas en un compacto (véase el teorema 5 en el p. 19.6).

Admitamos que existe una aplicación $f: A \to R^m$, $A \subset R^n$, continua en el compacto A, pero de manera no uniforme. En este caso existe tal $\varepsilon_0 > 0$ que para cualquier $\delta > 0$ existen unos puntos $x_\delta' \in A$ y $x_\delta'' \in A$, para los cuales tienen lugar las desigualdades

$$\rho(x'_\delta, x'_\delta) < \delta \quad \text{y} \quad \rho(f(x'_\delta), f(x'_\delta)) \geqslant \varepsilon_0.$$

Sea $\delta = \dfrac{1}{k}$, $x'^{(k)} \stackrel{\text{def}}{=} x'_{1/k}$, $x''^{(k)} \stackrel{\text{def}}{=} x''_{1/k}$, $k = 1, 2, \ldots$. Como A es un compacto, de la sucesión $[x'^{(k)}]$ se puede separar una subsucesión convergente $[x'^{(k_s)}]$ cuyo límite $x^{(0)}$ está contenido en el conjunto A: $\lim\limits_{s\to\infty} x'^{(k_s)} = x^{(0)} \in A$. En este caso de

$$\rho(x''^{(k_s)}, x^{(0)}) \leqslant \rho(x'^{(k_s)}, x''^{(k_s)}) + \rho(x'^{(k_s)}, x^{(0)}) < \frac{1}{k_s} + \rho(x'^{(k_s)}, x^{(0)}) \to 0,$$

cuando $s \to \infty$, se deduce que la subsucesión $[x''^{(k_s)}]$ de la segunda sucesión $[x''^{(k)}]$ también converge hacia el punto $x^{(0)}$.

Cabe señalar ahora que de la continuidad de la aplicación f en el punto $x^{(0)}$ se desprende que

$$\lim_{s\to\infty} f(x'^{(k_s)}) = \lim_{s\to\infty} f(x''^{(k_s)}) = f(x^{(0)}),$$

y puesto que

$$\rho(f(x''^{(k_s)}), f(x'^{(k_s)})) \leqslant \rho(f(x''^{(k_s)}), f(x^{(0)})) + \rho(f(x^{(0)}), f(x'^{(k_s)})) \to 0, \text{ para } s \to \infty,$$

entonces, $\lim\limits_{s\to\infty} \rho(f(x''^{(k_s)}), f(x'^{(k_s)})) = 0$. Esto contradice la condición $\rho(f(x''^{(k_s)}), f(x'^{(k_s)})) \geqslant \varepsilon_0$. □

Con ayuda de las propiedades demostradas de las aplicaciones continuas se puede obtener una propiedad de las regiones (es decir, de los conjuntos abiertos linealmente conexos, véase el p. 18.2) que será útil en lo que sigue. Enunciemos dicha propiedad también en forma de un lema.

Lema 5. *Un conjunto abierto constituye una región cuando, y sólo cuando, sus dos puntos cualesquiera pueden ser unidos mediante una línea quebrada íntegramente dispuesta en el conjunto.*

DEMOSTRACIÓN. La suficiencia de la condición enunciada no requiere demostración. En efecto, si en cierto conjunto abierto $G \subset R^n$ cualesquiera dos puntos pueden unirse mediante cierta quebrada, íntegramente dispuesta en el conjunto, entonces, puesto que toda quebrada es una curva (véase el p. 16.5), cualesquiera dos puntos del conjunto G resultan unidos en él por una curva lo que testimonia, según la definición (véase la definición 25 en el p. 18.2), que el conjunto abierto G es linealmente conexo, es decir, es una región (véase la definición 26 en el mismo punto).

Demostremos la necesidad de las condiciones del lema. Sea G una región del espacio R^n. Examinemos los puntos $x \in G$ e $y \in G$. Por definición de la región, existe una curva $\gamma = \{r(t), a \leqslant t \leqslant b\}$, que une en G los puntos x e y, es decir, $r(a) = x$, $r(b) = y$ y $r(t) \in G$, $a \leqslant t \leqslant b$. La curva γ representa en sí la imagen continua del segmento $[a, b]$ que es un compacto y, por lo tanto (véase el lema 3), la curva también es un compacto. Por cuanto el compacto γ y el conjunto cerrado $R^n \setminus G$ no se intersecan, la distancia entre ellos es superior a cero (véase el lema 7 en el p. 18.2). Por consiguiente, existe un número $\eta > 0$ tal que $\rho(\gamma, R^n \setminus G) > \eta$.

La aplicación $r(t)$, $a \leqslant t \leqslant b$, del segmento $[a, b]$, siendo continua, es también uniformemente continua (véase el lema 4). Por esto existe tal $\delta > 0$, que para cualesquiera dos puntos $t' \in [a, b]$ y $t'' \in [a, b]$, que satisfacen la condición $|t'' - t'| < \delta$, se verifica la desigualdad

$$\rho(r(t''), \quad r(t')) < \eta.$$

De aquí se deduce que para cualquier partición $\tau = \{t_i\}_{i=0}^{i=k}$ del segmento $[a, b]$ de finura $\delta_\tau < \delta$ todos los puntos de la quebrada λ_τ cos vértices $r(t_i)$, $i = 0, 1, \ldots, k$ estarán contenidos en G (¿por qué?). Por consiguiente, $\lambda_\tau \subset G$.

Puesto que el origen y el extremo de la quebrada λ_τ los consituyen el origen y el extremo, respectivamente, de la curva γ, es decir, los puntos prefijados arbitrariamente x e y de G, hemos demostrado, pues, que cualesquiera dos puntos de una región pueden unirse mediante una quebrada. □

Supongamos ahora que $X \subset R_x^n$, $D \subset R_y^m$ $y = f(x)$ es la aplicación del conjunto X en R_y^m con la particularidad de que $f(X) \subset D$ y $z = g(y)$ es la aplicación de D en R_z^p, es decir, $f: X \to D$, $g: D \to R_z^p$. En este caso tiene sentido la composición $g \circ f$: $X \to R_z^p$, que aplica el conjunto $X \subset R_x^n$ en el espacio p-dimensional R_z^p: $(g \circ f) \overset{\text{def}}{=} g(f(x))$, $x \in X$.

Hemos de notar que si la aplicación $f(x)$ del conjunto X es continua en el punto $x^{(0)} \in X$, y $g(y)$ está definida en cierto entorno del punto $y^{(0)} = f(x^{(0)})$, siempre existe tal entorno U_x del punto $x^{(0)}$ que en el conjunto $X \cap U_x$ tiene sentido la composición $g \circ f$. En efecto, sea U_y un entorno del punto $y^{(0)}$ en el que viene definida la aplicación $g(y)$; de acuerdo con la definición 3, para dicho entorno existe tal entorno U_x que $f(U_x \cap X) \subset U_y$. Es evidente que para todos los puntos $x \in U_x \cap X$ tiene sentido precisamente la composición $g \circ f$.

Recordemos, además, que, de conformidad con la terminología introducida para las funciones (véase el p. 1.2°), la aplicación $f: X \to R_y^m$, $X \subset R_x^n$ se llama *biunívoca* o bien *inyección* si a los puntos diferentes del conjunto X les corresponden en esta aplicación distintos puntos. En este caso suele decirse también que el conjunto X se aplica biunívocamente mediante dicha aplicación sobre el conjunto $f(X)$, es decir, $f: X \to f(X)$ es una biyección. Cumplida esta condición, en el conjunto $f(X)$ existe una aplicación inversa unívoca (una función inversa) $f^{-1}(y) = x$, donde x es tal que $f(x) = y$. Por ello, $f^{-1}[f(x)] = x$, es decir, ésta es una *aplicación idéntica* (se denomina aplicación idéntica del conjunto X aquella que a todo punto $x \in X$ le pone en correspondencia el mismo punto).

Definición 5. *Si la aplicación f del conjunto $X \subset R_x^n$ en el espacio R_y^m es biunívoca y continua sobre X y una aplicación inversa f^{-1} es continua sobre $f(X)$, entonces f se llama aplicación homeomorfa o homeomorfismo, y el conjunto $f(X)$, imagen homeomorfa del conjunto X, o bien, que es lo mismo, un conjunto homeomorfo respecto del conjunto X.*

Obviamente, si f es un homeomorfismo del conjunto X, f^{-1} es el homeomorfismo del conjunto $f(X)$.

Cuando se realiza la aplicación homeomorfa de un conjunto abierto sobre otro conjunto abierto, las imágenes de los subconjuntos abiertos son también abiertas. En efecto, si f es una aplicación homeomorfa del conjunto abierto G sobre otro conjunto abierto Γ, V es un subconjunto abierto del conjunto G, $W = f(V)$, enton-

ces $V = f^{-1}(W)$, es decir, V es una imagen del conjunto W, al realizarse la aplicación continua f^{-1} del conjunto abierto Γ, y, por consiguiente, W es una preimagen del conjunto abierto V en esta aplicación. Por ello, de acuerdo con el lema 2, el conjunto W es abierto.

Consideraremos ahora la composición de las aplicaciones continuas.

Lema 6. *Sea $f: X \to R^m$, $X \subset R^n$; $g: D \to R^s$, $D \supset f(X)$. Si la aplicación f es continua en el punto $x^{(0)} \in X$, y g es continua en el punto $f(x^{(0)})$, entonces la composición $g \circ f$ es también continua en el punto $x^{(0)}$.*

La demostración de dicha afirmación puede realizarse por un método análogo al empleado en la demostración del teorema 6, p. 5.16 y del teorema 2, p. 19.5. Este último método se basa en la definición de la continuidad en términos de los entornos. Con el fin de evitar la repitición, esta vez demostraremos el lema, partiendo de la definición de continuidad en términos de las sucesiones.

Sea $x^{(k)} \in X$, $k = 1, 2, \ldots$, y $\lim\limits_{k\to\infty} x^{(k)} = x^{(0)}$. En este caso $f(x^{(k)}) \in D$ y, por ser continua la aplicación f en el punto $x^{(0)}$, tenemos

$$\lim_{k\to\infty} f(x^{(k)}) = f(x^{(0)}). \tag{41.29}$$

En vista de que la aplicación g es continua en el punto $f(x^{(0)})$, para cualquier sucesión $y^{(k)} \in D$, $k = 1, 2, \ldots$, $\lim\limits_{k\to\infty} y^{(k)} = f(x^{(0)})$ tiene lugar $\lim\limits_{k\to\infty} g(y^{(k)}) = g(f(x^{(0)}))$. En particular, debido a (41.29), para $y^{(k)} = f(x^{(k)})$, se tiene

$$\lim_{k\to\infty} g(f(x^{(k)})) = g(f(x^{(0)})).$$

Esto precisamente implica la continuidad de la composición $g \circ f$ en el punto $x^{(0)}$. □

Ejercicio 7. Demuéstrese que la aplicación biunívoca continua de un compacto del espacio R^n en cierto espacio R^m es un homeomorfismo.

Como conclusión determinemos qué se entenderá por imagen de una curva, para una aplicación continua dada, y demostremos el lema sobre las imágenes continuas de los conjuntos linealmente conexos.

Sea f una aplicación continua del conjunto $X \subset R^n_x$ en el espacio R^m_y y Γ, una curva *dispuesta íntegramente en el conjunto* X, es decir, se ha dado una clase de las aplicaciones equivalentes de segmentos en el conjunto X (*véase* el § 16).

Sea

$$x(t), \quad a \leqslant t \leqslant b,$$

una de las representaciones de la curva Γ. Una curva en el espacio R^m_y cuya representación es la aplicación

$$f[x(t)], \quad a \leqslant t \leqslant b,$$

se denomina *imagen de la curva* Γ en la aplicación f y se designa por $f(\Gamma)$.

Esta definición es correcta, puesto que, con las suposiciones adoptadas, $f(x(t))$, $a \leqslant t \leqslant b$, es una aplicación continua de un segmento en el espacio y, por consiguiente, define cierta curva.

Lema 7. *Sea $f: X \to R^m$ una aplicación continua del conjunto linealmente conexo $X \subset R^n$ en el espacio R^m. En este caso el conjunto $f(X)$ es también linealmente conexo.*

En la forma más breve: *la imagen continua de un conjunto linealmente conexo es linealmente conexa.*

DEMOSTRACIÓN. Supongamos que X es un conjunto linealmente conexo y f es su aplicación continua en R^m. Para demostrar que el conjunto $f(X)$ es linealmente conexo, hay que demostrar que sus dos puntos cualesquiera pueden unirse en $f(X)$ mediante una curva continua (véase la definición 25 en el p. 18.2). Sea $y^{(1)} \in f(X)$ e $y^{(2)} \in f(X)$; elijamos dos puntos cualesquiera $x^{(1)} \in f^{-1}(y^{(1)})$ y $x^{(2)} \in f^{-1}(y^{(2)})$. Como $x^{(1)} \in X$, $x^{(2)} \in X$ y X es linealmente conexo, existe tal curva Γ, que su origen coincide con el punto $x^{(1)}$, el extremo coincide con el punto $x^{(2)}$ y todos los puntos de ella pertenecen al conjunto X.

La curva $f(\Gamma)$ es la curva buscada. En efecto, su origen es el punto $y^{(1)} = f(x^{(1)})$ y el extremo, el punto $y^{(2)} = f(x^{(2)})$. Todos los demás puntos de la curva pertenecen al conjunto $f(X)$. De este modo, $f(X)$ es un conjunto linealmente conexo. □

Ejercicios. 8. La aplicación $f: R^2 \to R^2$ viene dada de modo siguiente: $(x, y) \mapsto (2x, 3y)$. ¿En qué ésta transforma la circunferencia $x^2 + y^2 = 1$?

9. Hállese la imagen de la recta $x = 2$ del plano Oxy, realizándose la aplicación $f: R^2 \to R^2$, definida del modo siguiente: $(x, y) \mapsto (xy, y)$.

10. En el plano Oxy se ha dado una recta $x = c$ ($c = \text{const} \neq 0$). Hállese su imagen en la aplicación $f: R^2 \to R^2$ con las funciones coordenadas ($e^x \cos y$, $e^x \operatorname{sen} y$).

41.5. APLICACIONES VECTORIALES

Al estudiar las aplicaciones derivables (su definición se dará en el p. 41.7) resulta más conveniente que el espacio R^n, donde se dispone el conjunto que se aplica, y el espacio R^m, en el cual se realiza la aplicación, sean considerados como espacios euclídeos vectoriales (véase el p. 18.4). Para simplificar, un vector n-dimensional con coordenadas $(x_1, \ldots, x_n)$ se designará mediante el mismo símbolo x que se ha usado para designar un punto del espacio puntual n-dimensional con las mismas coordenadas. Esto no nos conducirá a una equivocación, pues tanto el punto del espacio n-dimensional como el vector n-dimensional representan un surtido ordenado de n números naturales.

Sea $X \subset R^n$ y $f: X \to R^m$, donde, en nuestro caso, la aplicación f pone en correspondencia a todo vector $x \in X$ cierto vector $y = f(x) \in R^m$. Las aplicaciones de este tipo se llamarán *vectoriales*.

Si $e_1, \ldots, e_n$ son los vectores coordenados en el espacio R^n (véase el p. 18.4), $\varepsilon_1, \ldots, \varepsilon_m$ son los vectores coordenados en el espacio R^m y $x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i$, $y = (y_1, \ldots, y_m) = \sum_{j=1}^{m} y_j \varepsilon_j$ e $y = f(x)$, entonces cualquier coordenada y_j, $j = 1, 2, \ldots, m$, del vector y es también una función del vector $x \in X$, y, por lo tanto, una función de sus coordenadas $x_1, \ldots, x_n$:

$$y_j = f_j(x) = f_j(x_1, \ldots, x_n), \quad j = 1, 2, \ldots, m: \tag{41.30}$$

Igual que en el caso de un espacio puntual (véase (41.26)), las funciones (41.30) se denominan *funciones coordenadas de la aplicación* f y se escribe $f = (f_1, \ldots, f_m)$

La interpretación de los puntos n-dimensionales $(x_1, \ldots, x_n)$ en calidad de vectores no impide, por supuesto, que se consideran tales propiedades de las aplicaciones como son su continuidad y continuidad uniforme. Por ello, todo lo dicho en el punto anterior sobre las aplicaciones queda en vigor para las aplicaciones vectoriales. Recordemos, además, que para la distancia $\rho(x, y)$ entre los vectores x e y es válida la fórmula (véase la vórmula (18.37) en el punto 18.4) $\rho(x, y) = |x - y|$.

A título de ejemplo señalemos que la longitud $|x|$ del vector $x \in R^n$ es una función continua en R^n. Esto proviene de la desigualdad (18.36): como para todo $x_0 \in R^n$ y todo $x \in R^n$ se verifica la desigualdad $||x| - |x_0|| \leqslant |x - x_0|$, entonces

$$\lim_{x \to x_0} |x| = \lim_{|x - x_0| \to 0} |x| = |x_0|.$$

41.6. APLICACIONES LINEALES

Consideraremos una clase especial de las aplicaciones del espacio R^n en R^m, denominadas *lineales*.

Definición 6. *Una aplicación $f: R^n \to R^m$ se llama lineal (o, en forma más completa, lineal homogénea), si para cualesquiera dos vectores $x' \in R^n$, $x'' \in R^n$ y dos números $\lambda' \in \boldsymbol{R}$, $\lambda'' \in \boldsymbol{R}$ se verifica la igualdad*

$$f(\lambda' x' + \lambda'' x'') = \lambda' f(x') + \lambda'' f(x'').$$

De esta definición se deduce por inducción que, realizándose la aplicación lineal f, cualquier combinación lineal finita de vectores $x^{(j)} \in R^n$ se aplica en la misma combinación lineal de imágenes $f(x^{(j)})$, $j = 1, 2, \ldots, k$, de dichos vectores

$$f\left(\sum_{j=1}^{k} \lambda_j x^{(j)}\right) = \sum_{j=1}^{k} \lambda_j f(x^{(j)}).$$

Las aplicaciones homogéneas lineales se llaman corrientemente *operadores lineales*. Del operador lineal $f: R^n \to R^m$ se dice que él actúa de R^n en R^m.

De la definición del operador lineal se desprende inmediatamente que la composición $g \circ f$ de operadores lineales $f: R^n \to R^m$ y $g: R^m \to R^s$ es también un operador lineal $g \circ f: R^n \to R^s$.

Sea $f: R^n \to R^m$ un operador lineal. La imagen de todo vector coordenado $e_j \in R^n$, $j = 1, 2, \ldots, n$, es, en la aplicación f, un vector del espacio R^m y, por ende, se descompone según los vectores coordenados $\varepsilon_i \in R^m$, $i = 1, 2, \ldots, m$. Designemos los coeficientes de esta descomposición con a_{ij}:

$$f(e_j) = \sum_{i=1}^{m} a_{ij} \varepsilon_i.$$

Sea $y = f(x)$, $x = \sum_{j=1}^{n} x_j e_j$ e

$$y = \sum_{i=1}^{m} y_i \varepsilon_i. \tag{41.31}$$

Por ser lineal la aplicación f, obtenemos

$$y = f(x) = f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j f(e_j) =$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\varepsilon_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\varepsilon_i. \qquad (41.32)$$

Al comparar los coeficientes de la descomposición del vector y según los vectores coordenados $\varepsilon_1, \ldots, \varepsilon_m$ en (41.31) y (41.32), obtendremos

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n. \end{aligned} \qquad (41.33)$$

Viceversa, es fácil comprobar que toda aplicación $f: R^n \to R^m$, cuyas funciones coordenadas tienen la forma (41.33), es un operador lineal.

La matriz

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & & \ldots\ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \qquad (41.34)$$

se llama *matriz del operador lineal* f.

Evidentemente, si (41.34) es una matriz del operator lineal f, para cualquier $x = \sum_{j=1}^{n} x_j e_j$ tiene lugar (véase (41.32)) la descomposición

$$f(x) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\varepsilon_i. \qquad (41.35)$$

Ejemplo. Sea π_i un operador de proyección sobre el i-ésimo eje coordenado, es decir,

$$\pi_i(y) = \pi_i(y_1, \ldots, y_m) = y_i \qquad (41.36)$$

(i es un número fijo entre los números $1, 2, \ldots, m$). En este caso π_i es un operador lineal con la matriz cuadrada de orden m compuesta sólo por ceros salvo el i-ésimo elemento de la diagonal principal el cual es igual a uno:

$$\begin{pmatrix} 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ & & \ldots\ldots\ldots & & & & \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ & & \ldots\ldots\ldots & & & & \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Con ayuda de los operadores de proyección π_i, $i = 1, 2, \ldots, n$, se establece fácilmente la relación entre una aplicación vectorial arbitraria $f: X \to R^m$, $X \subset R^n$ y sus funciones coordenadas f_i (véase (41.30)):

$$f_i = \pi_i \circ f, \tag{41.37}$$

es decir, toda función coordenada f_i, $i = 1, 2, \ldots, m$, es una composición de la aplicación f con el operador de proyección π_i.

Si $m = 1$, es decir, el operador lineal $f: R^n \to R$ aplica el espacio R^n en el conjunto de todos los números reales, se denomina corrientemente *funcional lineal*.

En virtud de (41.33), toda funcional lineal tiene por expresión

$$y = a_1x_1 + \ldots + a_nx_n, \tag{41.38}$$

donde $a_1, \ldots, a_n$ son ciertos números reales.

Al designar con a un vector de coordenadas $(a_1, \ldots, a_n)$, llegamos a que toda funcional lineal $f: R^n \to R$ tiene por expresión

$$f(x) = (a, x),$$

donde con (a, x) está designado el producto escalar de los vectores a y x. Es evidente también lo recíproco: toda aplicación de la forma $x \to (a, x)$ es una funcional lineal $f: R^n \to R$.

Ejercicio 11. Establézcase cuáles de las aplicaciones a seguir son lineales:

a) $f: R^3 \to R^2$, siendo $f(x, y, z) = (x, z)$;
b) $f: R^4 \to R^4$, siendo $f(x) = -x$, donde x es un vector arbitrario en R^4;
c) $f: R^3 \to R^3$, siendo $f(x) = x + (0, -1, 0)$ donde x es un vector arbitrario en R^3;
d) $f: R^2 \to R^2$, seindo $f(x, y) = (2x + y, y)$;
e) $f: R^2 \to R^2$, siendo $f(x, y) = 2x, y - x)$;
f) $f: R^2 \to R^2$, siendo $f(x, y) = (y, x)$;
g) $f: R^2 \to R$, siendo $f(x, y) = xy$.

Recordemos las definiciones de algunas operaciones sobre matrices (conocidas del álgebra). Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices rectangulares con el mismo número de columnas y filas, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, la *suma* de ellas se determina como una matriz cuyo elemento c_{ij} es la suma de los elementos correspondientes de las matrices A y B, es decir,

$$c_{ij} \stackrel{\text{def}}{=} a_{ij} + b_{ij}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.$$

Se llama *producto de la matriz A* por el número λ una matriz cuyos elementos c_{ij} se obtienen todos de los elementos correspondientes de la matriz A multiplicándolos por λ

$$c_{ij} \stackrel{\text{def}}{=} \lambda a_{ij}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.$$

Si el número de las columnas en la matriz $A = (a_{ij})$ es igual al número de las filas de la matriz $B = (b_{jk})$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, s$, entonces el *producto* AB de las matrices A y B se determina como una matriz compuesta de los elementos c_{ik}, que se determinan según las fórmulas:

$$c_{ik} \stackrel{\text{def}}{=} \sum_{j=1}^{n} a_{ij}b_{jk}, \quad i = 1, 2, \ldots, m, \quad k = 1, 2, \ldots, s.$$

He aquí dos propiedades de los operadores lineales las cuales nos harán falta en lo sucesivo.

1°. *Si f y g son operadores lineales, $f: R^n \to R^m$, $g: R^n \to R^m$ y λ y μ, números arbitrarios, entonces $\lambda f + \mu g$ es también un operador lineal que actúa de R^n en R^m, con la particularidad de que si A y B son las matrices de los operadores lineales f y g, la suma $\lambda A + \mu B$ será la matriz del operador $\lambda f + \mu g$.*

La demostración de esta afirmación se efectúa por comprobación inmediata: si

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \ldots, m,$$

son funciones coordenadas de la aplicación f, mientras que

$$z_i = \sum_{j=1}^{n} b_{ij}x_j, \quad i = 1, 2, \ldots, m,$$

son funciones coordenadas de la aplicación g, entonces para las funciones coordenadas de la aplicación $\lambda f + \mu g$ tendremos (al sumar y multiplicar por números los vectores, sus coordenadas se suman y se multiplican por los mismos números)

$$\lambda y_i + \mu z_i = \lambda \sum_{j=1}^{n} a_{ij}x_j + \mu \sum_{j=1}^{n} b_{ij}x_j = \sum_{j=1}^{n} (\lambda a_{ij} + \mu b_{ij})x_j,$$

es decir, primero, las funciones coordenadas de la aplicación $\lambda f + \mu g$ son funciones lineales y, segundo, los elementos c_{ij} de la matriz de la aplicación $\lambda f + \mu g$ los constituyen los números $c_{ij} = \lambda a_{ij} + \mu b_{ij}$, es decir, los elementos de la matriz $\lambda A + \mu B$, donde $A = (a_{ij})$, $B = (b_{ij})$. □

2°. *Si f y g son operadores lineales, $f: R^n \to R^m$, $g: R^m \to R^s$, su composición $g \circ f$ es también un operador lineal $R^n \to R^s$ y la matriz de la composición es igual al producto de matrices de las aplicaciones g y f.*

Realicemos nuevamente la comprobación inmediata de la afirmación. Si

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \ldots, m,$$

son funciones coordenadas de la aplicación f y

$$z_k = \sum_{i=1}^{m} b_{ki}y_i, \quad k = 1, 2, \ldots, s,$$

son funciones coordenadas de la aplicación g, entonces

$$z_k = \sum_{i=1}^{m} b_{ki}y_i = \sum_{i=1}^{m} b_{ki} \sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} b_{ki}a_{ij} \right) x_j,$$

es decir, primero, las funciones coordenadas de la composición $g \circ f$ son funciones

lineales, y, segundo, los elementos c_{kj} de la matriz de la composición se obtienen de los elementos de las matrices a_{ij} y b_{ki} de los operadores f y g según la regla

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}. \tag{41.39}$$

Según se ha dicho, tal matriz (c_{kj}) se llama precisamente producto de las matrices (b_{ki}) y (a_{ij}). □

Observemos que todo operador lineal $f: R^n \to R^m$ es una aplicación continua del espacio R^n, dado que todas sus funciones coordenadas (41.33) son continuas, pues son lineales.

La longitud del vector $x \in R^n$, como se ha notado en el p. 41.5, es una función continua en el espacio R^n. Por esta razón, si $f: R^n \to R^m$ es un operador lineal, entonces la función $|f(x)|$, representando una composición de dos funciones continuas, será también continua en R^n.

Por cuanto la bola unidad $Q^n = \{x \in R^n : |x| \leqslant 1\}$ es un compacto, para todo operador lineal $f: R^n \to R^m$ la restricción de la función continua $|f|: R^n \to R$ a la bola Q^n, es decir, la función $|f|: Q^n \to R$, está acotada:

$$\sup_{|x| \leqslant 1} |f(x)| < +\infty. \tag{41.40}$$

Definición 7. *Para el operador lineal (en particular, para una funcional lineal, cuando $m = 1$) $f: R^n \to R^m$ el número* $\sup_{|x| \leqslant 1} |f(x)|$ *lleva el nombre de norma**) *del operador y se denota con $\|f\|$:*

$$\|f\| \stackrel{\text{def}}{=} \sup_{|x| \leqslant 1} |f(x)|. \tag{41.41}$$

En virtud de la desigualdad (41.40), la norma de cualquier operador lineal es finita.

Estimemos la longitud de la imagen del vector $x \in R^n$ en términos de la norma del operador f y la longitud x del propio vector. Para todo $x \neq 0$, $x \in R^n$, el vector $\xi = \dfrac{x}{|x|}$ tiene la longitud $1: |\xi| = \left|\dfrac{x}{|x|}\right| = \dfrac{1}{|x|}|x| = 1$. Por ello, al utilizar la linealidad del operador f, la propiedad (18.34) de la longitud del vector y la definición (41.41), obtendremos

$$|f(x)| = \left|f\left(|x|\frac{x}{|x|}\right)\right| = \left||x| f\left(\frac{x}{|x|}\right)\right| = |x|\left|f\left(\frac{x}{|x|}\right)\right| \leqslant$$

$$\leqslant |x| \sup_{|\xi| \leqslant 1} |f(\xi)| = |x| \, \|f\|,$$

es decir,

$$|f(x)| \leqslant \|f\| \, |x|. \tag{41.42}$$

---

*) La definición general de la norma se dará en el p. 57.3.

De esta desigualdad se infiere que para $|x| < 1$ es válida la desigualdad $|f(x)| < \|f\|$. Recordemos que una función, continua en el compacto, alcanza en éste su valor máximo (véase el teorema 3 en el p. 19.6). Por eso una función $|f|: Q^n \to R$, siendo continua en el compacto Q^n, alcanza en éste su valor máximo

$$\|f\| \sup_{|x| \leqslant 1} |f(x)| = \max_{|x| \leqslant 1} |f(x)|,$$

y, como para $|x| < 1$ tiene lugar la desigualdad $|f(x)| < \|f\|$, el máximo mencionado se alcanza cuando $|x| = 1$, es decir, en la esfera unidad $S^{n-1} = \{x : |x| = 1\}$. De este modo

$$\|f\| = \max_{|x| = 1} |f(x)|. \tag{41.43}$$

Demos a conocer una expresión más para la norma de un operador lineal

$$\|f\| = \sup_{x \in R^n,\ x \neq 0} \frac{|f(x)|}{|x|}. \tag{41.44}$$

Demostrémosla. Empleando otra vez la propiedad de la longitud (18.34) de un vector, la linealidad de la aplicación f y la fórmula (41.41), obtenemos

$$\sup_{x \in R^n,\ x \neq 0} \frac{|f(x)|}{|x|} = \sup_{x \in R^n,\ x \neq 0} \left| \frac{1}{|x|} f(x) \right| =$$

$$= \sup_{x \in R^n,\ x \neq 0} \left| f\left(\frac{x}{|x|}\right) \right| = \sup_{|\xi| = 1} |f(\xi)| = \|f\|.$$

No es difícil estimar la norma $\|f\|$ del operador lineal f en términos de los elementos de su matriz (41.34). Fijándonos en que el cuadrado de longitud del vector es igual a la suma de cuadrados de sus coordenadas y aplicando la fórmula (41.35) y la desigualdad de Cauchy-Schwarz (18.2), tendremos

$$|f(x)|^2 = \left| \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \varepsilon_i \right|^2 =$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right)^2 \leqslant \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) =$$

$$= \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right) |x|^2.$$

De aquí, para todo $x \neq 0$, $x \in R^n$:

$$\frac{|f(x)|}{|x|} \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

Por ello, en virtud de (41.44),

$$\|f\| = \sup_{x \in R^n,\ x \neq 0} \frac{|f(x)|}{|x|} \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} \tag{41.45}$$

De este modo queda demostrada la desigualdad (41.40), esta vez, de "modo algebraico".

41.7. APLICACIONES DERIVABLES

Pasemos ahora a la definición de aplicaciones vectoriales derivables. Recordemos previamente que una función de n variables $f: X \to R$, $X \subset R^n$*), definida en el entorno del punto $x = (x_1, \ldots, x_n) \in X$, se llama *derivable* en este punto, si existen tales constantes $a_1, \ldots, a_n$ (son derivadas parciales de la función f en este punto:

$$a_i = \frac{\partial f}{\partial x_i}(x)), \text{ que}$$

$$f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) = a_1 h_1 + \ldots + a_n h_n + o(h),\ h \to 0, \tag{41.46}$$

donde $h = (h_1, \ldots, h_n)$.

La aplicación lineal (funcional lineal)

$$(h_1, \ldots, h_n) \mapsto a_1 h_1 + \ldots + a_n h_n$$

en la fórmula (41.46) se denomina *diferencial de la función* f en el punto x. Al designarla mediante $D(x)$, obtenemos

$$D(x)(h) = a_1 h_1 + \ldots + a_n h_n.$$

De este modo, la definición de la derivabilidad (41.46) puede representarse en la forma

$$f(x + h) = f(x) + D(x)(h) + o(h),\quad h \to 0.$$

De modo análogo se determina, en el caso general, la derivabilidad de la aplicación.

Para las aplicaciones que se realizan de un espacio n-dimensional a un espacio m-dimensional "o pequeño" se determina de la manera siguiente: sea U un entorno del punto $x_0 \in X$, $\alpha: U \to R^m$; diremos que $\alpha = o(x)$ para $x \to x_0$, siempre que $|\alpha| = o(|x|)$, $x \to x_0$, es decir, si existe tal función $\varepsilon: U \to R$, que

$$|\alpha(x)| = \varepsilon(x)|x|, \quad x \in U \text{ y } \lim_{x \to x_0} \varepsilon(x) = 0. \tag{41.47}$$

Sin ninguna duda $|\alpha(x)|$ es la longitud de un vector en el espacio R^m y $|x|$ es la longitud del vector en el espacio R^n. Para las expresiones del tipo $o(x)$, que son vecto-

*) Con R se designa, como siempre, el conjunto de todos los números reales.

res, quedan vigentes las reglas habituales de operaciones con el símbolo "o pequeño", por ejemplo, $o(x) + o(x) = o(x)$ para $x \to x_0$, etc.

Definición 8. *Sea U un entorno del punto $x \in R^n$ en el espacio R^n. Un aplicación $f: U \to R^m$ se llama derivable en el punto x, si existe tal aplicación lineal (operador lineal) $l: R^n \to R^m$ que*

$$f(x + h) = f(x) + l(h) + o(h), \quad h \to 0, h \in R^n. \tag{41.48}$$

El operador lineal l se llama diferencial de la aplicación f en el punto x y se denota con $D(x)$ o, más detalladamente, con $D_f(x)$.

Haciendo uso de esta designación, la definición de la derivabilidad (41.48) puede escribirse en la forma

$$f(x + h) = f(x) + D_f(x)(h) + o(h), \quad h \to 0. \tag{41.49}$$

La matriz de la diferencial $D_f(x)$ (véase (41.34)) *se llama derivada de la aplicación f en el punto x y se designa mediante $f'(x)$.*

Observemos que de la fórmula (41.48) proviene inmediatamente que una aplicación, derivable en el punto x, es continua en él:

$$\lim_{h \to 0} f(x + h) = f(x).$$

Teorema 3. *Si la aplicación $f: X \to R^m$, $X \subset R^n$, es derivable en el punto $x \in X$, entonces su diferencial en este punto se define unívocamente.*

Corolario. *La diferencial de una aplicación lineal coincide con la misma aplicación.*

DEMOSTRACIÓN DEL TEOREMA. Supongamos que a la par con la igualdad (41.48) se cumple también la igualdad

$$f(x + h) = f(x) + l_1(h) + o(h), \quad h \to 0, \tag{41.50}$$

donde $l_1: R^n \to R^m$, l_1 es un operador lineal. Restando una de estas igualdades de la otra, obtenemos

$$l(h) - l_1(h) = o(h) \quad \text{para} \quad h \to 0,$$

es decir, existe tal función $\varepsilon(h)$, definida en cierto entorno V del origen de coordenadas del espacio R^n, es decir, $\varepsilon: V \to R$, que $\lim_{h \to 0} \varepsilon(h) = 0$ y que para todo $h \in V$ tiene lugar la igualdad

$$|l(h) - l_1(h)| = \varepsilon(h)\,|h|. \tag{41.51}$$

Tomemos ahora arbitrariamente $k \in R^n$; para todo t suficientemente pequeño tendremos $tk \in V$. Por eso, en (41.51) para tales t podemos poner $h = tk$:

$$|l(tk) - l_1(tk)| = \varepsilon(tk)\,|tk|.$$

Como $|tk| = |t|\,|k|$ y las aplicaciones l y l_1 son lineales, tendremos

$$l(tk) - l_1(tk) = t[l(k) - l_1(k)],$$

y, por ende,

$$|l(k) - l_1(k)| = \varepsilon(tk)\,|k|. \tag{41.52}$$

Pero $\lim_{t \to 0} tk = 0$, por lo cual, en virtud de la propiedad de la función ε, tenemos también $\lim_{t \to 0} \varepsilon(tk) = 0$. Pasando al límite para $t \to 0$ en (41.52), obtendremos $|l(k) - l_1(k)| = 0$, es decir, para cualquier $k \in R^n$

$$l(k) = l_1(k).$$

Esto precisamente indica que $l = l_1$. □

DEMOSTRACIÓN DEL COROLARIO. Sea $f: R^n \to R^m$ un operador lineal. Por ser lineal para cualesquiera $x \in R^n$ y $h \in R^n$, tenemos

$$f(x + h) = f(x) + f(h),$$

es decir, la igualdad (41.48) se cumple para $l = f$ y $o(h) = 0$. Debido a la unicidad de la diferencial, $D_f(x) = f$. □

Teorema 4 ((linealidad de la diferencial). *Si las aplicaciones $f: \mathbf{X} \to R^m$ y g: $X \to R^m$, $X \subset R^n$, son derivables en el punto $x \in X$, entonces para cualesquiera números λ y μ la combinación lineal $\lambda f + \mu g$ será también derivable en el punto x y*

$$D_{\lambda f + \mu g}(x) = \lambda D_f(x) + \mu D_g(x).$$

DEMOSTRACIÓN. Debido a la derivabilidad de las aplicaciones f y g en el punto x, tenemos (véase (41.49)):

$$f(x + h) = f(x) + D_f(x)(h) + o(h), \quad h \to 0$$
$$g(x + h) = g(x) + D_g(x)(h) + o(h), \quad h \to 0;$$

de aquí

$$(\lambda f(x + h) + \mu g(x + h)) =$$
$$= [\lambda f(x) + \mu g(x)] + [\lambda D_f(x) + \mu D_g(x)](h) + o(h), \quad h \to 0.$$

Como $\lambda D_f(x) + \mu D_g(x)$ es una aplicación lineal (véase el p. 41.6), entonces, de conformidad con la definición 8, la aplicación lineal $\lambda D_f(x) + \mu D_g(x)$ es la diferencial de la aplicación $\lambda f + \mu g$. □

Teorema 5. *Supongamos que $X \subset R^n$, $D \subset R^m$, $f: X \to D$, $g: D \to R^r$, con la particularidad de que la aplicación f es derivable en el punto $x \in X$ y g, en el punto $f(x)$. En este caso la composición $g \circ f$ es derivable en el punto x y su diferencial en el mismo es igual a la composición de diferenciales de las aplicaciones f y g:*

$$D_{g \circ f}(x) = D_g(f(x)) \circ D_f(x). \tag{41.53}$$

Corolario. *Cumplidas las condiciones del teorema, la derivada de la composición de aplicaciones es igual al producto de las derivadas:*

$$(g \circ f)'(x) = g'(f(x))\, f'(x). \tag{41.54}$$

Como se ve de las fórmulas aducidas, gracias a la elección adecuada de las definiciones y anotaciones, en las formulaciones de los teoremas tiene lugar la analogía completa con el caso unidimensional.

DEMOSTRACIÓN. Por ser derivable la aplicación f, tenemos

$$(g \circ f)(x + h) = g(f(x + h)) = g(f(x) + D_f(x)(h) + o(h)), \quad h \to 0. \tag{41.55}$$

De este modo, el argumento de la función g en el punto $y = f(x)$ ha recibido un incremento

$$k = D_f(x)(h) + o(h). \tag{41.56}$$

Por ello, de (41.55) tenemos en virtud de que la función g es derivable:

$$(g \circ f)(x + h) = g(y + k) = g(y) + D_g(y)(k) + o(k), \quad k \to 0. \tag{41.57}$$

Dado que (véase la desigualdad (41.42)).

$$|D_f(x)(h)| \leqslant \|D_f(x)\| \, |h|, \tag{41.58}$$

donde la norma $\|D_f(x)\|$ del operador lineal $D_f(x)$ es un número no negativo, entonces para la función $k = k(h)$, definida por la igualdad (41.56), obtendremos

$$\lim_{h \to 0} k = 0. \tag{41.59}$$

Más aún, resulta justa la estimación

$$|k| \leqslant \|D_f(x)\| \, |h| + |o(h)|, \quad h \to 0,$$

y como, para h suficientemente pequeños, tiene lugar la desigualdad $|o(h)| < |h|$, para h de este género es lícita también la estimación

$$|k| \leqslant (\|D_f(x)\| + 1) \, |h|. \tag{41.60}$$

Luego, de la definición de $o(k)$ (véase (41.47)) se desprende que existe una función $\varepsilon(k)$ tal que

$$\lim_{k \to 0} \varepsilon(k) = 0 \tag{41.61}$$

y $|o(k)| = \varepsilon(k) \, |k|$. Por esta razón, en virtud de (41.60), para h mencionados suficientemente pequeños, se verifica la desigualdad

$$|o(k)| = \varepsilon(k) \, |k| \leqslant \varepsilon(k)(\|D_f(x)\| + 1) \, |h|. \tag{41.62}$$

Por cuanto de (41.59) y (41.61) se infiere que $\lim_{h \to 0} \varepsilon(k) = 0$, se tiene, por consiguiente,

$$\varepsilon(k)(\|D_f(x)\| + 1) \, |h| = o(h) \quad \text{para} \quad h \to 0,$$

entonces de (41.62) tenemos $|o(k)| \leqslant o(h)$, $h \to 0$, de donde

$$o(k) = o(h) \quad \text{para} \quad h \to 0.$$

Esto significa que la fórmula (41.57) puede escribirse en la forma

$$(g \circ f)(x + h) = g(y) + D_g(y)(k) + o(h), \quad h \to 0, \tag{41.63}$$

donde k se determina por la fórmula (41.56).

Consideremos ahora el sumando medio en el segundo miembro de la igualdad (41.63). Puesto que la aplicación $D_g(y)$ es lineal, tenemos

$$D_g(y)(k) = D_g(y)(D_f(x)(h) + o(h)) = \\ = D_g(y)(D_f(x)(h)) + D_g(y)o(h), \quad h \to 0. \tag{41.64}$$

Debido a la desigualdad (41.42), tendremos $|D_g(y)o(h)| \leqslant \|D_g(y)\| \, |o(h)|$, por lo cual

$$D_g(y)o(h) = o(h), \quad h \to 0;$$

por consiguiente, de (41.64) obtendremos:

$$D_g(y)(k) = D_g(y)(D_f(x)(h)) + o(h) = (D_g(y) \circ D_f(x))(h) + o(h), \quad h \to 0.$$

Sustituyendo la expresión obtenida para $D_g(y)$ en (41.63) y tomando en consideración que $y = f(x)$, tendremos en definitiva

$$(g \circ f)(x + h) = g(f(x)) + (D_g(f(x)) \circ D_f(x))(h) + o(h), \quad h \to 0.$$

Como la composición de operadores lineales es un operador lineal, entonces, por la unicidad de la diferencial, el operador $D_g(f(x)) \circ D_f(x)$ es la diferencial de la composición $f \circ g$, es decir, la fórmula (41.53) queda demostrada.

La fórmula (41.54) se deduce de ésta en seguida, puesto que en la composición de los operadores lineales sus matrices se multiplican. □

Teorema 6. *La aplicación $f = (f_1, \ldots, f_m)$: $X \to R^m$, $X \subset R^n$, es derivable en el punto $x \in X$, si, y sólo si, todas las funciones coordenadas suyas f_i: $X \to R$, $i = 1, 2, \ldots, m$, son derivables en dicho punto. En este caso los elementos a_{ij} de la matriz de la diferencial $D_f(x)$ son derivadas parciales correspondientes de las funciones coordenadas:*

$$a_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \; i = 1, 2, \ldots, m, \; j = 1, 2, \ldots, n.$$

En otras palabras, la derivada $f'(x)$ es la matriz de Jacobi del sistema de funciones f_i (véase la definición 2 del p. 41.3),

$$f'(x) = \begin{pmatrix} \dfrac{\partial f_1(x)}{\partial x_1} & \ldots & \dfrac{\partial f_1(x)}{\partial x_n} \\ \ldots\ldots\ldots\ldots & & \ldots\ldots\ldots\ldots \\ \dfrac{\partial f_m(x)}{\partial x_1} & \ldots & \dfrac{\partial f_m(x)}{\partial x_n} \end{pmatrix} \tag{41.65}$$

y se llama también *matriz de Jacobi de la aplicación f* en el punto x.

DEMOSTRACIÓN 1. Las funciones coordenadas $f_i = \pi_i \circ f$ (véase (41.37)) son una composición de dos aplicaciones derivables: la aplicación f, que es derivable en el punto f por hipótesis, y la proyección π_i (véase (41.36)) que es derivable en todo el espacio R^m como cualquier operador lineal $R^m \to R$. Por consiguiente, de acuerdo con el teorema 5, las funciones f_i, $i = 1, 2, \ldots, m$, son derivables en el punto x.

2. Supongamos que todas las funciones coordenadas $f_i = \pi_i \circ f$ de la aplicación son derivables en el punto x. Teniendo presente (41.46), esto implica que existen tales constantes a_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ que

$$f_i(x + h) = f_i(x) + a_{i1}h_1 + \ldots + a_{in}h_n + o(h), \; h \to 0, \; i = 1, 2, \ldots, m. \tag{41.66}$$

De aquí, como se sabe (véase (20.2)), proviene que los coeficientes a_{ij} de los incrementos h_j que adquieren los argumentos x_j son las derivadas parciales correspondientes de las funciones f_i:

$$a_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \; i = 1, 2, \ldots, m, \; j = 1, 2, \ldots, n. \tag{41.67}$$

Designemos mediante $l\colon R^n \to R^m$ un operador lineal con la matriz (a_{ij}). Por cuanto $(o(h), \ldots, o(h)) = o(h)$*), las igualdades (41.66) pueden escribirse en la forma

$$f(x + h) = f(x) + l(h) + o(h), \quad h \to 0.$$

Esto precisamente significa la derivabilidad de la aplicación f, además de (41.67) se infiere la validez de la fórmula (41.65). □

OBSERVACIÓN 1. En vista de las fórmulas (41.54) y (41.65), el corolario del teorema 5 significa que la matriz de Jacobi de la composición de aplicaciones f y g es igual al producto de las matrices de Jacobi de las aplicaciones mencionadas.

Esto, por otra parte, se deduce también directamente de la fórmula de derivación de una función compuesta: si $z_k = g_k(y_1, \ldots, y_m)$, $k = 1, 2, \ldots, s$, e $y_i = f_i(x_1, \ldots, x_n)$, $i = 1, 2, \ldots, m$, entonces (véase (20.26))

$$\frac{\partial z_k}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial z_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}, \; k = 1, 2, \ldots, s, \; j = 1, 2, \ldots, n,$$

que, de acuerdo con la regla de multiplicación de las matrices (véase el p. 41.6) significa que la matriz $\left(\frac{\partial z_k}{\partial x_j}\right)$ es un producto de las matrices $\left(\frac{\partial z_k}{\partial y_i}\right)$ y $\left(\frac{\partial y_i}{\partial x_j}\right)$:

$$\left(\frac{\partial z_k}{\partial x_j}\right) = \left(\frac{\partial z_k}{\partial y_i}\right)\left(\frac{\partial y_i}{\partial x_j}\right).$$

Definición 9. *En el caso cuando $m = n$, el determinante*

$$\det\left(\frac{\partial f_i(x)}{\partial x_j}\right)$$

de la matriz de Jacobi (41.65) *se llama determinante de Jacobi o jacobiano de la aplicación $f\colon X \to R^n$, $X \subset R^n$, en el punto $x \in X$ y se denota* (véase el p. 41.3)

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \quad \text{o} \quad \frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}.$$

OBSERVACIÓN 2. Por el curso de álgebra conocemos que al multiplicar las matrices cuadradas, sus determinantes se multiplican, razón por la cual, cumplidas las condiciones del teorema 5 en el caso de $m = n = s$, el jacobiano de la composición de las aplicaciones f y g es igual al producto de jacobianos de las aplicaciones f y g

---

*) La notación $(o(h), \ldots, o(h)) = o(h)$ significa que el vector cuyas coordenadas son infinitésimos de orden superior a h, es de por sí un infinitésimo de orden más elevado que h cuando $h \to 0$. La coincidencia que en el caso dado tiene lugar entre las designaciones del vector y de sus coordenadas se debe a que para el vector n-dimensional se ha elegido, para no complicar la notación, la designación x en la que no está reflejada la dimensión del vector. Ésta se pone de manifiesto sólo cuando el vector se escribe con ayuda de las coordenadas $x = (x_1, \ldots, x_n)$.

$$\frac{\partial(z_1, \dots, z_n)}{\partial(x_1, \dots, x_n)} = \frac{\partial(z_1, \dots, z_n)}{\partial(y_1, \dots, y_n)} \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}. \tag{41.68}$$

En efecto,

$$\frac{\partial(z_1, \dots, z_n)}{\partial(x_1, \dots, x_n)} = \det\left(\frac{\partial z_k}{\partial x_j}\right) = \det\left(\frac{\partial z_k}{\partial y_i}\right)\left(\frac{\partial y_i}{\partial x_j}\right) =$$

$$= \det\left(\frac{\partial z_k}{\partial y_i}\right) \det\left(\frac{\partial y_i}{\partial x_j}\right) = \frac{\partial(z_1, \dots, z_n)}{\partial(y_1, \dots, y_n)} \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}.$$

OBSERVACIÓN 3. Supongamos que $X \subset R^n$ e Id: $X \to X$ es una aplicación idéntica del conjunto X sobre sí mismo. En la forma coordenada se escribe como la condición de igualdad de las cordenadas de los puntos correspondientes en esta aplicación a la imagen y la preimagen, es decir, las funciones coordenadas tienen por expresión

$$f_i(x) = x_i, \; i = 1, 2, \dots, n, \; (x_1, \dots, x_n) \in X.$$

Si $x^{(0)}$ es un punto interior del conjunto X, estas funciones pueden derivarse en dicho punto y, puesto que $\frac{\partial x_i}{\partial x_j} = 0$ para $i \neq j$ y $\frac{\partial x_i}{\partial x_i} = 1$, entonces la matriz de Jacobi de la aplicación idéntica es una matriz unidad

$$X = \begin{pmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1 \end{pmatrix}.$$

Supongamos ahora que $U \subset R^n$, $V \subset R^n$ y f: $U \to V$ es una aplicación biunívoca (inyectiva), mientras que f^{-1}: $f(U) \to U$ es la aplicación inversa a la primera. En este caso, para todo punto $x \in U$ tenemos $f^{-1}(f(x)) = x$, es decir, la composición $f^{-1} \circ f$ es una aplicación idéntica.

Supongamos que la aplicación f es derivable en el punto $x_0 \in U$ (por tanto, x_0 es un punto interior del conjunto U, pues solo para tales puntos viene determinada la noción de derivabilidad), mientras que la aplicación inversa f^{-1} es derivable en el punto $f(x_0)$. Como $f^{-1} \circ f$ es un aplicación idéntica, en vista de la fórmula (41.54) tenemos

$$(f^{-1})'f' = (f^{-1} \circ f)' = (\mathrm{Id})' = X. \tag{41.69}$$

Pasando de esta igualdad de las matrices a los jacobianos de éstas, obtendremos

$$\det(f^{-1})' \det f' = 1, \tag{41.70}$$

ya que $\det X = 1$.

Si la aplicación f viene dada por las funciones coordenadas (41.30), la fórmula (41.70) puede escribirse en la forma

$$\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} = 1. \tag{41.71}$$

De esta fórmula se deduce que con las suposiciones admitidas tanto el jacobiano de la aplicación f en el punto x, como el de la aplicación inversa f^{-1} en el punto $f(x)$ no se anulan.

Escribamos la fórmula (41.71) en otra forma

$$\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} = \frac{1}{\dfrac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}}. \tag{41.72}$$

Dicha fórmula representa una generalización evidente de la fórmula para la derivada de la función inversa de una sola variable: $\dfrac{dx}{dy} = \dfrac{1}{\dfrac{dy}{dx}}$.

Como conclusión enunciaremos dos definiciones útiles.

Definición 10. *Una aplicación $f: X \to R^m$, $X \subset R^n$, derivable en todo punto $x \in X$, se denomina aplicación derivable del conjunto X.*

Obviamente, si una aplicación es derivable en el conjunto X, entonces, cualquiera que sea el punto $x \in X$, de acuerdo con la definición 8, la aplicación f queda definida en cierto entorno del punto, es decir, X es un conjunto abierto.

De conformidad con el teorema 6, la aplicación $f = (f_1, \dots, f_n)$ es derivable en el conjunto X cuando, y sólo cuando, son derivables en dicho conjunto todas las funciones coordenadas suyas $f_1, \dots, f_n$. Si todas las funciones coordenadas son continuamente derivables en X, es decir, si todas sus primeras derivadas parciales son continuas en X, entonces la aplicación f se llama *aplicación continuamente derivable del conjunto X.*

Definición 11. *Una aplicación homeomorfa $f: G \to D$, donde G y D son conjuntos abiertos del espacio R^n, se llama difeomorfa o difeomorfismo, si son derivables tanto la propia aplicación citada como la aplicación inversa $f^{-1}: D \to G$.*

41.8. APLICACIONES CON UN JACOBIANO DISTINTO DE CERO. PRINCIPIO DE CONSERVACIÓN DE LA REGIÓN

Consideraremos ante todo la cuestión acerca de la existencia de una aplicación inversa a la dada. Como se sabe, en el caso de $n = 1$, para una función continuamente derivable en cierto segmento, la condición de que la derivada de ésta no se reduzca a cero (la que conlleva a su monotonía estricta) resulta suficiente para que exista una función unívoca continuamente derivable que sea inversa respecto de la dada. Cuando n es arbitrario, el caso se complica considerablemente: las correspondientes condiciones puntuales impuestas sobre las propiedades de derivabilidad de la aplicación sólo permiten afirmar que la aplicación inversa existe localmente, es decir, en un entorno del punto. Con más precisión, queda válido el siguiente teorema.

Teorema 7. *Sea*

$$y = f(x) = \begin{cases} y_1 = f_1(x_1, \dots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ y_n = f_n(x_1, \dots, x_n) \end{cases} \tag{41.73}$$

una aplicación continuamente derivable del conjunto abierto $G \subset R^n$ en el espacio R^n. Si el jacobiano de esta aplicación no se anula en el punto $x^{(0)} \in G$, existen, pues, tales entornos U_x y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)} = f(x^{(0)})$, que $f(x)$, $x \in U_x$, es la aplicación biunívoca del entorno U_x sobre el entorno U_y, y la aplicación inversa es continuamente derivable en el conjunto U_y.

Corolario. *Sea f una aplicación continuamente derivable del conjunto abierto $G \subset R^n$ en el espacio R^n. Si el jacobiano de la aplicación f no es nulo en G, la imagen del conjunto G en esta aplicación es también un conjunto abierto.*

DEMOSTRACIÓN. Examinemos las funciones

$$F_i(x, y) = f_i(x_1, \dots, x_n) - y_i, \quad i = 1, 2, \dots, n.$$

Están definidas para todo $y = (y_1, \dots, y_n) \in R_y^n$ y para todo $x = (x_1, \dots, x_n) \in G \subset R_x^n$. Con su ayuda el sistema de ecuaciones (41.73) que prefijan la aplicación f, se escribirá en la forma

$$F_i(x, y) = 0, \quad i = 1, 2, \dots, n. \tag{41.74}$$

Las funciones $F_i(x, y)$ están definidas y son continuamente derivables en cierto entorno del punto $(x^{(0)}, y^{(0)})$ (como tal entorno puede ser tomado, por ejemplo, $G \times R_y^n$),

$$F_i(x^{(0)}, y^{(0)}) = 0 \text{ y } \frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)}\bigg|_{(x^{(0)}, y^{(0)})} = \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}\bigg|_{x^{(0)}} \neq 0.$$

De este modo, resultan cumplidas todas las condiciones del teorema 2 de este párrafo sobre la resolubilidad del sistema de ecuaciones.

En virtud de este teorema, las ecuaciones (41.74), o, que es lo mismo, el sistema (41.73) pueden ser resueltos y, además, de un modo único, respecto de las variables $x_1, \dots, x_n$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$. Más detalladamente esto significa que existen tales entornos U_x^* y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$, $x^{(0)} \in U_x^*$, $y^{(0)} \in U_y$, y tal aplicación única

$$x = g(y) = \begin{cases} x_1 = g_1(y_1, \dots, y_n), \\ \dots\dots\dots\dots\dots\dots \\ x_n = g_n(y_1, \dots, y_n) \end{cases} \tag{41.75}$$

que aplica el entorno U_y en el entorno U_x^*, que para todo $y \in U_y$ tiene lugar la identidad

$$f[g(y)] = y.$$

En otras palabras, para todo punto $y \in U_y$ existe, y además, el único punto $x = g(y) \in U_x^*$ el que, al realizarse la aplicación f, pasa al punto y. De este modo, $x \in f^{-1}(y) \cap U_x^*$ y $g(y)$ es una aplicación unívoca, continuamente derivable e inversa respecto de f en U_y: $g = f^{-1}$.

Pongamos $U_x = U_x^* \cap f^{-1}(U_y)$. Entonces, U_x es un conjunto abierto, pues representa una intersección de dos conjuntos abiertos U_x^* y $f^{-1}(U_y)$ (el carácter abierto del conjunto $f^{-1}(U_y)$ se desprende de que es una preimagen del conjunto abierto U_y en la aplicación continua f, véase el lema 2 en el p. 41.4). Es evidente que

U_x se aplica biunívocamente sobre U_y y, dado que $x^{(0)} \in U_x^*$ y $f(x^{(0)}) = y^{(0)} \in U_y$, entonces $x^{(0)} \in U_x$, es decir, U_x es el entorno buscado del punto $x^{(0)}$. □

OBSERVACIÓN 1. Los entornos U_x y U_y que figuran en las condiciones del teorema 4 poseen, además, la propiedad adicional de que el jacobiano de la aplicación f del entorno U_x sobre el entorno U_y no se anula en el entorno U_x, mientras que el jacobiano de la aplicación inversa f^{-1} no se anula en el entorno U_y. Esto proviene inmediatamente de la fórmula (41.71). En efecto, en vista de que la aplicación f transforma biunívocamente el entorno U_x en el U_y y que f y f^{-1} son continuamente derivables, la fórmula mencionada puede emplearse para el caso de la aplicación f, considerada sobre el conjunto U_x. Conforme a esta fórmula, un producto de los jacobianos de las aplicaciones f y f^{-1} es igual a uno y, por lo tanto, cada uno de ellos no es igual a cero.

DEMOSTRACIÓN DEL COROLARIO. Supongamos que $y = f(x)$ es una aplicación continuamente derivable del conjunto abierto G en el espacio R^n y que $y^{(0)}$ es un punto arbitrario del conjunto $f(G)$. Elijamos un punto $x^{(0)}$ en la preimagen del punto $y^{(0)}$: $x^{(0)} \in f^{-1}(y^{(0)})$. Por consiguiente, $f(x^{(0)}) = y^{(0)}$. En virtud del teorema 4, existen tales entornos $U_x \subset G$ y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$ que $f(U_x) = U_y$. Por consiguiente, $U_y \subset f(G)$. En otras palabras, para todo punto $y_{(0)} \in f(G)$ existe su entorno contenido en el conjunto $f(G)$. De este modo, cualquier punto del conjunto $f(G)$ es interior para este conjunto lo que es un indicio de que $f(g)$ es un conjunto abierto. □

OBSERVACIÓN 2. Si en cierta aplicación f para los puntos $x^{(0)}$ e $y^{(0)} = f(x^{(0)})$ existen unos entornos respectivos U_x y U_y que se aplican biunívocamente por f uno sobre el otro, suele decirse que *la aplicación f es localmente biunívoca en el punto* $x^{(0)}$.

Si en este caso la aplicación f es continua en U_x y f^{-1} es continua en U_y, entonces f se llama *aplicación localmente homeomorfa* en el punto $x^{(0)}$ u *homeomorfismo local*. Por fin, si el homeomorfismo local citado es un difeomorfismo, la aplicación en consideración recibe el nombre de *difeomorfismo local* en el punto dado (véanse las definiciones de homeomorfismo y difeomorfismo en el p. 41.4 y el p. 41.7).

Empleando la terminología mencionada, podemos decir que la aplicación f, considerada en el teorema 4, es una aplicación localmente difeomorfa en todo punto en el que el jacobiano es destinto de cero.

Teorema 8 (principio de conservación de la región). *La imagen de una región n-dimensional de un espacio n-dimensional en la aplicación continuamente derivada con un jacobiano que no se reduce a cero, es una región.*

DEMOSTRACIÓN. Sea G una región, $G \subset R^n$, e $y = f(x)$ una aplicación de G en R^n que satisface las condiciones del teorema. De acuerdo con el corolario del teorema 4, el conjunto $f(G)$ es abierto y, según el lema 7 del p. 41.4, es linealmente conexo. Por ello, si G es una región, el conjunto $f(G)$ será también una región, siempre que se cumplen las condiciones del teorema. □

Ejercicio 12. Constrúyase un ejemplo de una aplicación continuamente derivable de cierta región plana cuyo jacobiano nunca se reduce a cero y no es biunívoca.

41.9. FUNCIONES IMPLÍCITAS DEFINIDAS POR UNA ECUACIÓN EN LA QUE SE TRASTORNAN LAS CONDICIONES DE UNICIDAD. PUNTOS SINGULARES DE LAS CURVAS PLANAS

Ya sabemos que si las coordenadas de cierto punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ satisfacen la ecuación

$$F(x_1, \ldots, x_n) = 0 \tag{41.76}$$

y si en este punto la derivada $\dfrac{\partial F}{\partial x_i}$ no es nula, entonces bajo ciertas condiciones correspondientes impuestas sobre la continuidad de la misma función F y de la derivada citada, la ecuación (41.76) resulta resoluble en cierto entorno del punto $x^{(0)}$ respecto de x_i y la solución es una función continuamente derivable de las demás coordenadas.

Surge, naturalmente, la cuestión: ¿qué sucederá, cuando en el punto $x^{(0)}$ las derivadas parciales respecto de todos los argumentos se reduzcan a cero: definirá o no, en este caso, la ecuación (41.76) algunas funciones? Detendrémonos en esta cuestión, limitándonos a la consideración de un caso bidimensional, teniendo en cuenta la enorme dificultad con la que se resuelve.

Así pues, consideraremos una ecuación

$$F(x, y) = 0, \tag{41.77}$$

donde la función F está definida y es continuamente derivable en cierto entorno del punto (x_0, y_0) de tal género que

$$F(x_0, y_0) = 0. \tag{41.78}$$

Sea

$$F_x(x_0, y_0) = F_y(x_0, y_0) = 0 \tag{41.79}$$

Mostremos que incluso si se cumplen estas condiciones, la ecuación (41.77) se resuelve, a veces, en el entorno del punto (x_0, y_0) respecto de una de las variables, de suerte que se obtendrá una función continuamente derivable; no obstante, en el caso general, esto se puede hacer de una manera que no es única. De tal modo, la condición

$$F_x^2(x_0, y_0) + F_y^2(x_0, y_0) \neq 0, \tag{41.80}$$

que en nuestro caso no se cumple (véase (41.79)) y que permite aplicar el teorema 1 sobre las funciones implícitas a una de las variables, puede llamarse, naturalmente, condición de solubilidad unívoca de la ecuación (41.77).

Definición 12. *Un punto (x_0, y_0) cuyas coordenadas satisfacen las condiciones* (41.78) *y* (41.79) *se denomina punto singular de la ecuación* (41.77).

Un punto singular se llama aislado, si existe un entorno suyo en el cual el punto citado es el único punto singular.

En el lenguaje geométrico esto significa que si la ecuación (41.77) es una representación implícita de alguna curva, entonces en el entorno de los puntos singulares de esta ecuación la curva no representa, en el caso general, la gráfica de cierta función unívoca suave (lo que tiene lugar, si se cumplen las condiciones (41.80)); aquí

pueden surgir diferentes singularidades que ahora serán el objeto de nuestra consideración.

Introduzcamos, para abreviar, las siguientes designaciones

$$F_{xx}(x_0, y_0) = F^0_{xx}, \quad F_{xy}(x_0, y_0) = F^0_{xy}, \quad F_{yy}(x_0, y_0) = F^0_{yy}.$$

Teorema 9. *Supongamos que la función $F(x, y)$ está definida y es dos veces continuamente derivable en cierto entorno de un punto singular aislado (x_0, y_0) de la ecuación* (41.77) *y sea*

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} \neq 0.$$

En este caso, si

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} > 0, \tag{41.81}$$

entonces (x_0, y_0) es una solución aislada de la ecuación (41.77), *es decir, existe un entorno del punto (x_0, y_0) tal que ningún punto de él, salvo (x_0, y_0) satisface la ecuación* (41.77); *en cambio, si*

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} < 0, \tag{41.82}$$

la ecuación (41.77) *es resoluble en cierto entorno del punto (x_0, y_0) pero no unívocamente: hay dos diferentes funciones derivables que satisfacen la ecuación* (41.77). *Por ello (x_0, y_0) se llama en este caso punto doble.*

Por ejemplo, si

$$F^0_{yy} \neq 0, \tag{41.83}$$

existen dos funciones derivables $f_1(x)$ y $f_2(x)$, definidas en cierto entorno del punto x_0, tales que en dicho entorno $F(x, f_1(x)) \equiv 0$, $F(x, f_2(x)) \equiv 0$, con la particularidad de que $f_1(x_0) = f_2(x_0) = y_0$, mientras que las derivadas de las funciones $f_1(x)$ y $f_2(x)$ en el punto x_0 son diferentes raíces de la ecuación

$$F^0_{xx} + 2F^0_{xy}k + F^0_{yy}k^2 = 0^{*)}. \tag{41.84}$$

DEMOSTRACIÓN. Supongamos que se cumple la condición (41.81). Junto con (41.79) es suficiente para que exista en el punto (x_0, y_0) un extremo estricto de la función $F(x, y)$ (véase el teorema 3 en el p. 40.2). Por eso existe un entorno U del punto (x_0, y_0) tal que para $(x, y) \in U$ y $(x, y) \neq (x_0, y_0)$ o bien siempre $F(x, y) > F(x_0, y_0)$, o bien siempre $F(x, y) < F(x_0, y_0)$ y como $F(x_0, y_0) = 0$, se tiene $F(x, y) \neq 0$ para cualesquiera $(x, y) \in U$, $(x, y) \neq (x_0, y_0)$, es decir, (x_0, y_0) es la solución aislada de la ecuación (41.77)**).

Supongamos ahora que se cumple la condición (41.82). Desarrollemos la función $F(x, y)$, rigiéndonos por la fórmula de Taylor, en el entorno del punto (x_0, y_0)

---

*) Las raíces de esta ecuación son reales y diferentes, en virtud de las condiciones (41.82) y (41.83).

**) Para demostrar esta afirmación se usa no el hecho de que (x_0, y_0) es el punto singular aislado, sino que es sólo un punto simplemente singular en el que se cumple la condición (41.81).

hasta los sumandos de segundo orden. Tomando en consideración las condiciones (41.78) y (41.79), obtendremos

$$F(x, y) = \frac{1}{2}[F^0_{xx}(x - x_0)^2 + 2F^0_{xy}(x - x_0)(y - y_0) + F^0_{yy}(y - y_0)^2] + o(r^2), \tag{41.85}$$

donde $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$. Pongamos $x - x_0 = r\cos\varphi$, $y - y_0 = r\operatorname{sen}\varphi$. Evidentemente, (r, φ) son las coordenadas polares del punto (x, y), con la particularidad de que a título de origen del sistema polar de coordenadas se ha adoptado el punto (x_0, y_0).

En estas coordenadas

$$F(x, y) = \frac{r^2}{2}(F^0_{xx}\cos^2\varphi + 2F^0_{xy}\cos\varphi\operatorname{sen}\varphi + F^0_{yy}\operatorname{sen}^2\varphi) + o(r^2) =$$

$$= \frac{r^2}{2}P(\varphi) + o(r^2), \tag{41.86}$$

donde

$$P(\varphi) = F^0_{xx}\cos^2\varphi + 2F^0_{xy}\cos\varphi\operatorname{sen}\varphi + F^0_{yy}\operatorname{sen}^2\varphi, \tag{41.87}$$

o bien, para $\varphi \neq \frac{\pi}{2}(2k + 1)$, $k = 0, \pm 1, \pm 2, \ldots$,

$$P(\varphi) = \cos^2\varphi(F^0_{xx} + F^0_{xy}\operatorname{tg}\varphi + F^0_{yy}\operatorname{tg}^2\varphi). \tag{41.88}$$

Supongamos ahora que queda cumplida también la condición (41.83). Sean k_1 y k_2 las raíces de la ecuación (41.84) y sean $\varphi_1 = \operatorname{arctg} k_1$ y $\varphi_2 = \operatorname{arctg} k_2$. En este caso

$$\varphi_1 \neq \pm\pi/2, \quad \varphi_2 \neq \pm\pi/2, \tag{41.89}$$

y de (41.88) se infiere que

$$P(\varphi) = \cos^2\varphi(\operatorname{tg}\varphi - \operatorname{tg}\varphi_1)(\operatorname{tg}\varphi - \operatorname{tg}\varphi_2). \tag{41.90}$$

De la fórmula (41.90) se ve que la función $P(\varphi)$ se anula, para $\varphi \neq \frac{\pi}{2}(2k + 1)$, $k = 0, \pm 1, \pm 2, \ldots$, cuando $\varphi = \varphi_1 + k\pi$ y $\varphi = \varphi_2 + k\pi$, $k = 0, \pm 1, \pm 2, \ldots$, con la particularidad de que al pasar el argumento por estos valores, la función cambia de signo. Será cómodo para nosotros interpretar $P(\varphi)$ como función de un punto de la circunferencia C de radio igual a la unidad (este radio se elige a fin de simplificar el caso, para que las longitudes de los arcos coincidan con los ángulos φ) y centro en el punto (x_0, y_0).

Sea $\varepsilon > 0$. Designemos mediante $U_1 = U_1(\varepsilon)$ un ángulo abierto determinado por la desigualdad $\varphi_1 - \varepsilon < \varphi < \varphi_1 + \varepsilon$, es decir,

$$U_1 = \{(r, \varphi): \varphi_1 - \varepsilon < \varphi < \varphi_1 + \varepsilon\},$$

y pongamos correspondientemente

$$U_2 = \{(r, \varphi): \varphi_2 - \varepsilon < \varphi < \varphi_2 + \varepsilon\};$$

eligiendo, además, $\varepsilon > 0$ tan pequeño que U_1 y U_2 no se corten y no contengan en sí los semiejes de las ordenadas y, por consiguiente, las semirrectas verticales en general (lo último siempre puede conseguirse en vista de las condiciones (41.89)).

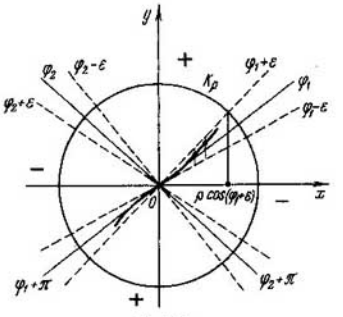


Fig. 166

Sean U_1^* y U_2^* los ángulos centrales simétricos con U_1 y U_2 respecto del punto (x_0, y_0):

$$U_1^* = \{(r, \varphi): \varphi_1 + \pi - \varepsilon < \varphi < \varphi_1 + \pi + \varepsilon\},$$
$$U_2^* = \{(r, \varphi): \varphi_2 + \pi - \varepsilon < \varphi < \varphi_2 + \pi + \varepsilon\}.$$

Por la elección del número ε los conjuntos U_1, U_2, U_1^* y U_2^* no se cortan dos a dos (fig. 166).

Examinemos ahora $P(\varphi)$ como función de un punto de la circunferencia mencionada C. Para simplificar, designemos también mediante φ un punto de la circunferencia C, al cual corresonde el ángulo polar φ. Eliminemos en la circunferencia C los intervalos con centros en los puntos φ_1, φ_2, $\varphi_1 + \pi$ y $\varphi_2 + \pi$ de longitud 2ε*); por la elección adecuada de $\varepsilon > 0$, estos intervalos no tienen puntos comunes. El conjunto restante, que se designará con B, es limitado y cerrado y, por lo tanto, representa un compacto. En B la función $P(\varphi)$ es continua y no se anula, a consecuencia de lo cual

$$\inf_{\varphi \in B} |P(\varphi)| = \mu > 0. \tag{41.91}$$

Designemos mediante K_ρ un círculo cerrado con centro en el punto (x_0, y_0) y radio ρ:

$$K_\rho = \{(r, \varphi): 0 \leqslant r \leqslant \rho\},$$

y mediante L_ρ, un conjunto que se obtiene por sustracción (en el sentido de la teoría de conjuntos, véase el p. 1.1) de los conjuntos U_1, U_2, U_1^* y U_2^* del círculo K_ρ. Es ob-

*) Se llama intervalo de longitud 2ε en la circunferencia con centro en un punto, cuyo ángulo polar es igual a φ_0, un conjunto de sus puntos cuyos ángulos polares φ satisfacen la desigualdad $\varphi_0 - \varepsilon < \varphi < \varphi_0 + \varepsilon$.

vio que debido a (41.91) se tiene

$$\inf_{(r,\,\varphi)\,\in\,L_\rho} |P(\varphi)| = \mu > 0.$$

Ahora, observando que de (41.86) se deduce

$$F(x, y) = \frac{r^2}{2}[P(\varphi) + \alpha(r, \varphi)], \tag{41.92}$$

donde $\lim_{r \to 0} \alpha(r, \varphi) = 0$, elijamos $\rho > 0$ de modo tal que para $r \leqslant \rho$ se verifique la desigualdad

$$|\alpha(r, \varphi)| < \mu. \tag{41.93}$$

Entonces, de (41.92) se desprende que para todos los puntos $(r, \varphi) \in L_\rho$ la expresión que figura en el segundo miembro de la fórmula (41.92) es del mismo signo que $P(\varphi)$.

El conjunto L_ρ se compone de cuatro sectores cerrados (véase fig. 166), en cada uno de los cuales, a excepción de su centro, la función $P(\varphi)$ y, por ende, debido a la elección adecuada de ρ, la función $F(x, y)$ adquieren valores de un mismo signo, y en los sectores vecinos, de signos contrarios.

Examinemos ahora el ángulo $U_1 = U_1(\varepsilon)$. Sea, para concretar, $0 \leqslant \varphi_1 < \pi/2$. La intersección de la clausura $\overline{U}_1$ del ángulo U_1 con la recta vertical $x = x^*$, $x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon)$, representa un segmento en cuyos extremos superior e inferior la función $F(x^*, y)$ toma los valores de signos diferentes. La función $F(x^*, y)$ considerada como función de una sola variable y con x^* fijado, es continua en el segmento mencionado y por eso se anula en cierto punto y^* del segmento, es decir, para todo x^*, donde $x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon)$, existe por lo menos un punto y^* tal que

$$F(x^*, y^*) = 0, \quad (x^*, y^*) \in U_1(\varepsilon) \cap K_\rho. \tag{41.94}$$

Definamos $y = f_1(x)$ como una función que pone en correspondencia al número x^* el número y^*:

$$f_1(x^*) = y^*, \quad x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon).$$

Mostremos que para ε y ρ suficientemente pequeños la función f_1 está definida unívocamente, es decir, existen tales $\varepsilon > 0$ y $\rho > 0$ que con x^* dado las condiciones (41.94) definen unívocamente y^*. Admitamos lo contrario. Tomemos las sucesiones $\varepsilon_n \to 0$ y $\rho_n \to 0$, para $n \to \infty$. En este caso existen dos sucesiones de los puntos que tienen abscisas iguales x_n y ordenadas distintas y'_n e y''_n, tales que

$$(x_n, y'_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}, \quad F(x_n, y'_n) = 0,$$
$$(x_n, y''_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}, \quad F(x_n, y''_n) = 0.$$

En virtud del teorema de Rolle, en el segmento $[y'_n, y''_n]$ de la recta $x = x_n$ existe tal punto y_n que

$$F_y(x_n, y_n) = 0. \tag{41.95}$$

siendo, además, en este caso $(x_n, y_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}$; por hipótesis (véase (41.79)) teníamos

$$F_y(x_0, y_0) = 0. \tag{41.96}$$

Según la fórmula de incrementos finitos aplicada a la función $F_y(x, y)$ se tiene

$$F_y(x_n, y_n) - F_y(x_0, y_0) = F_{yx}(\xi_n, \eta_n)(x_n - x_0) + F_{yy}(\xi_n, \eta_n)(y_n - y_0),$$
$$(\xi_n, \eta_n) \in U_1(\varepsilon_n) \cap K_{\rho_n},$$

de donde, en virtud de (41.95) y (41.96),

$$F_{xy}(\xi_n, \eta_n) + F_{yy}(\xi_n, \eta_n)\frac{y_n - y_0}{x_n - x_0} = 0. \tag{41.97}$$

Sea $(x_n, y_n) = (r_n, \psi_n)$. Evidentemente, $|\psi_n - \varphi_1| < \varepsilon_n$; y por esta razón de la condición $\varepsilon_n \to 0$ se deduce que $\psi_n \to \varphi_1$ cuando $n \to \infty$, y como $\operatorname{tg} \psi_n = \dfrac{y_n - y_0}{x_n - x_0}$, entonces

$$\lim_{n \to \infty} \frac{y_n - y_0}{x_n - x_0} = \operatorname{tg} \varphi_1 = k_1. \tag{41.98}$$

Pasando al límite en la igualdad (41.97) para $n \to \infty$, en vista de (41.98) tenemos

$$F^0_{xy} + F^0_{yy} k_1 = 0, \quad \text{es decir,} \quad k_1 = -\frac{F^0_{xy}}{F^0_{yy}};$$

sustituyendo este valor de la raíz en la ecuación (41.84), obtenemos

$$F^0_{xx} F^0_{yy} - F^{02}_{xy} = 0,$$

que contradice la condición (41.82).

Así pues, la función $y = f_1(x)$ se define realmente de modo unívoco cuando ε y ρ son suficientemente pequeños. En lo que sigue supondremos que ε y ρ se han elegido precisamente de la manera indicada.

Definamos adicionalmente la función f_1 en el punto x_0, al poner $y_0 = f_1(x_0)$. Obviamente, por la propia definición de la función $f_1(x)$, tenemos

$$F(x, f_1(x)) = 0, \quad x_0 \leqslant x \leqslant x_0 + \rho \cos(\varphi_1 + \varepsilon).$$

Mostremos que en el punto x_0 la función $f_1(x)$ tiene una derivada a la derecha y que esta derivada es igual a k_1. Sea $\varepsilon > 0$ arbitrariamente fijado. De lo dicho más arriba se desprende la existencia de tal $\rho = \rho(\varepsilon) > 0$ que la parte correspondiente de la gráfica de la función $f_1(x)$ se dispone íntegramente en $U_1(\varepsilon) \cap K_\rho$:

$$(x, f_1(x)) \in U_1(\varepsilon) \cap K_\rho, \quad x_0 \leqslant x \leqslant x_0 + \rho \cos(\varphi_1 + \varepsilon) \tag{41.99}$$

Tomemos $\delta = \rho \cos(\varphi_1 + \varepsilon)$ y sea x tal que $0 < x - x_0 < \delta$, $y = f_1(x)$ y $(x, y) = (r, \varphi)$. En virtud de (41.99), tenemos $|\varphi - \varphi_1| < \varepsilon$. Esto significa que $\lim\limits_{x \to x_0 + 0} \varphi = \varphi_1$, y, por lo tanto, $\lim\limits_{x \to x_0 + 0} \operatorname{tg} \varphi = \operatorname{tg} \varphi_1$. Puesto que $\operatorname{tg}\varphi = \dfrac{y - y_0}{x - x_0}$, de lo demostrado se deduce que

$$\lim_{x \to x_0 + 0} \frac{f_1(x) - f_1(x_0)}{x - x_0} = \lim_{x \to x_0 + 0} \frac{y - y_0}{x - x_0} = \operatorname{tg} \varphi_1,$$

es decir, la función $f_1(x)$ tiene derivada a la derecha en el punto x_0 y esta derivada es igual a $\operatorname{tg} \varphi_1 = k_1$.

De esta misma manera se demuestra, analizando el comportamiento de la función $F(x, y)$ en el ángulo U_1^*, que en el segmento $[x_0 - \delta', x_0]$ existe, para cierto $\delta' > 0$, tal función $f_1(x)$ que siendo $x_0 - \delta' \leqslant x \leqslant x_0$, se tiene

$$F(x, f_1(x)) \equiv 0, \quad (x, f_1(x)) \in U^*_1, \quad f_1(x_0) = k_1$$

(por derivada se entiende en el caso dado la derivada a la izquierda).

Si ρ se toma tan pequeño que en un entorno circular de radio ρ del punto (x_0, y_0) no se contengan otros puntos singulares de la ecuación (41.77), a excepción de (x_0, y_0), la función $f_1(x)$ será derivable en todo punto $x \neq x_0$. Esto se deduce en seguida del teorema sobre las funciones implícitas que fue demostrado anteriormente (véase el teorema 1 en el p. 41.1). De resultas hemos obtenido la función $f_1(x)$ que está definida en cierto entorno del punto x_0 y que posee todas las propiedades requeridas.

Análogamente se demuestra la existencia de la función $f_2(x)$ que también es una solución de la ecuación (41.77) y satisface las condiciones del teorema, con la particularidad de que la gráfica de esta función pasa en los ángulos U_2 y U_2^* y por el punto (x_0, y_0).

Si $F^0_{yy} = 0$ y $F^0_{xx} \neq 0$, entonces todos los razonamientos quedan los mismos; sólo conviene cambiar de lugar los ejes Ox y Oy, de suerte que obtendremos, como resultado, las soluciones de la ecuación (41.77) en forma de las funciones de la variable y: $f_1(y)$ y $f_2(y)$.

Si, por fin, $F^0_{xx} = F^0_{yy} = 0$ y, por lo tanto, $F^0_{xy} \neq 0$, entonces resulta más conveniente realizar el cambio de variables: $x = \xi + \eta$ e $y = \xi - \eta$ (gírense los ejes de coordenadas al ángulo $\pi/4$). En este caso (de lo que es fácil convencerse por derivabilidad directa)

$$F^0_{\xi\xi} = -F^0_{\eta\eta} = 2F^0_{xy} \neq 0, \; F^0_{\xi\eta} = 0,$$

es decir, en el nuevo sistema de coordenadas se obtendrá el caso ya estudiado. En particular, la ecuación (41.84) para los coeficientes angulares de las tangentes en el punto singular tiene, en el sistema de coordenadas ξ, η, la forma siguiente

$$k^2 - 1 = 0,$$

y, por consiguiente, $k_{1,2} = \pm 1$. En otras palabras, las bisectrices de los ángulos coordenados, que representan los ejes en el sistema antiguo de coordenadas x, y, son las tangentes a las gráficas de dos funciones definidas por la ecuación (41.77) en cierto entorno del punto singular que se considera. □

Si la ecuación $F(x, y) = 0$ es una representación implícita de alguna curva, en el punto singular (x_0, y_0) de esta ecuación la curva puede tener (aunque no sea necesariamente) ciertas singularidades, es decir, en el entorno del punto singular de dicha ecuación la curva no es, en el caso general, la gráfica de cierta función unívoca suave.

Cabe recordar también que un conjunto de los puntos cuyas coordenadas satisfacen la ecuación (41.77) no es (en el caso general) siempre una curva en el sentido de la definición de la curva dada paramétricamente, enunciada anteriormente (véase el p. 16.2*).

Ejemplos. 1. Sea dada una ecuación $y^2(x^2 + y^2 + 1) = 0$. Aquí, $F(x, y) = y^2(x^2 + y^2 + 1)$, por lo cual $F_x = 2xy^2$, $F_y = 2x^2y + 4y^3 + 2y$. Las condiciones para la presencia de un punto singular (41.78) y (41.79) nos dan en este caso

$$x_0 = 0, y_0 = 0.$$

De este modo, de punto singular sirve el punto (0, 0). No obstante, en este punto la curva, definida por la ecuación, no tiene singularidad, puesto que la ecuación (el factor $x^2 + y^2 + 1$ nunca se reduce a cero) es equivalente a la $y = 0$ y la curva en consideración representa una gráfica de la función explícita $y = f(x) \equiv 0$. Observemos que en el punto (0, 0) para este caso

$$F_{xx}F_{yy} - F_{xy}^2 = 0, \tag{41.100}$$

de lo que podemos convencernos con facilidad.

2. Para la ecuación

$$(x^2 + y^2)(x^2 + y^2 - 1) = 0, \tag{41.101}$$

las condiciones (41.79) se convierten en el siguiente sistema de ecuaciones:

$$2x^3 + 2xy^2 - x = 0,$$
$$2y^3 + 2x^2y - y = 0.$$

Al sumar y restar estas ecuaciones, obtenemos el sistema

$$(x + y)(2x^2 + 2y^2 - 1) = 0$$
$$(x - y)(2x^2 + 2y^2 - 1) = 0.$$

De aquí, o bien $x = y = 0$, o bien $2x^2 + 2y^2 - 1 = 0$, sin embargo, el punto (x, y) cuyas coordenadas satisfacen la última correlación no es una raíz de la ecuación (41.101) (para este punto $x^2 + y^2 = \frac{1}{2}$, y, por lo tanto, en el primer miembro no hay factor que se reduzca a cero).

De este modo, el único punto singular es (0, 0). Es fácil comprobar que aquí se cumple la condición (41.81) y, por ende, el punto (0, 0) es una raíz aislada de la ecuación (41.101). Desde el punto de vista geométrico, como se ve en seguida, la ecuación (41.101) define la circunferencia unidad y el centro de ésta (0, 0) (este conjunto no porta, evidentemente, ninguna curva dada paramétricamente en el sentido del p. 16.2°).

3. Para la ecuación

$$x^3 + y^3 - 3axy = 0, \tag{41.102}$$

las condiciones (41.79) para la existencia de un punto singular conducen al sistema de ecuaciones

$$x^2 - ay = 0,$$
$$y^2 - ax = 0,$$

de donde o bien $x = y = 0$, y en este caso dicho punto satisface la ecuación (41.102), o bien $x = a, y = a$, pero las coordenadas de este punto no constituyen la solución de la ecuación (41.102). Aquí, otra vez, (0, 0) es el único punto singular.

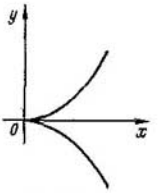


Fig. 167

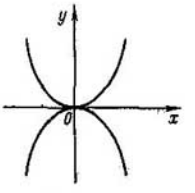


Fig. 168

No es difícil convencerse de que en el caso dado se cumplen las condiciones (41.82) y, por consiguiente, (0, 0) es un punto doble.

Geométricamente, para una curva cuya representación explícita es la ecuación (41.102) (se llama folio de Descartes y con ella ya chocamos en el p. 14.5) el punto (0, 0) es un *punto múltiple* (véase la fig. 70 en el tomo I).

4. Para la ecuación

$$y^2 - x^3 = 0, \tag{41.103}$$

(0, 0) es un punto singular; en dicho punto ya se cumple la condición (41.100) y, consecuentemente, en este caso no se cumplen las condiciones del teorema 6. Geométricamente, una curva expresada por la ecuación (41.103) y llamada parábola semicúbica $y = \pm x^{3/2}$, tiene en el punto (0, 0) una tangente y se dispone en el entorno de este punto por un lado respecto de la normal.

Los puntos de esta índole se llaman *puntos de retroceso* (fig. 167).

5. Para la ecuación

$$y^2 - x^4 = 0 \tag{41.104}$$

el punto (0, 0) es también singular y aquí se cumple nuevamente la condición (41.100). La ecuación (41.104) se descompone, evidentemente, en dos ecuaciones: $y = x^2$ e $y = -x^2$ que definen dos parábolas y estas últimas tienen en el punto (0, 0) una tangente común.

Los puntos singulares, en cierto entorno de los cuales la ecuación (41.77) define dos curvas continuamente derivables que tienen en el punto (x_0, y_0) una tangente común, se llaman *puntos autoadherentes* (fig. 168) de estas dos curvas.

Puede ocurrir que, cumplida la condición (41.100), el punto singular resulte ser la solución aislada de la ecuación (41.77) o su punto doble.

Como conclusión, ofrecemos algunas explicaciones para la ecuación (41.84). Si (x_0, y_0) es un punto singular de la ecuación (41.77), entonces, realizado el traslado paralelo del origen de coordenadas al punto (x_0, y_0), la ecuación (41.77) adquiere la forma

$$F^0_{xx}x^2 + 2F^0_{xy}xy + F^0_{yy}y^2 + o(x^2 + y^2) = 0, \tag{41.105}$$

(aquí, con x e y están designadas las coordenadas del punto en el nuevo sistema de coordenadas y por índice 0, puesto por arriba, se designan los valores de las deriva-

das parciales en el punto (0,0) de este sistema), de donde, con una exactitud de hasta los infinitésimos de orden superior, nuestra ecuación puede escribirse del modo siguiente:

$$F''_{xx}x^2 + 2F''_{xy}xy + F''_{yy}y^2 = 0. \tag{41.106}$$

Si se cumple la condición (41.82), el primer miembro de la ecuación (41.106) se descompone en dos factores reales, cada uno de los cuales, siendo igualado a cero, proporciona las tangentes a las dos ramas de la curva en el punto (0, 0) (véase (41.84)). Cuando se cumple la condición (41.81), el primer miembro de la ecuación (41.106) se descompone en dos factores complejos: "las tangentes son imaginarias". Esto es natural, puesto que no hay sentido en hablar aquí de una tangente, pues en el caso dado el punto singular es aislado.

Esta observación es de comodidad especial cuando se trata de determinar el carácter del punto singular en el caso de una curva algebraica, es decir, de la curva definida por la ecuación

$$P(x, y) = 0, \tag{41.107}$$

donde $P(x, y)$ es un polinomio de dos variables x e y. Si (0, 0) es un punto singular de esta ecuación, de las condiciones (41.78) y (41.79) se deduce que dicho polinomio no contiene término independiente ni tampoco los términos de primer orden, es decir, la ecuación (41.107) tiene por expresión

$$ax^2 + 2bxy + cy^2 + Q(x, y) = 0,$$

donde $Q(x, y)$ es un polinomio, todos los términos del cual son por lo menos de tercer orden. El comportamiento de las soluciones de esta ecuación se determina por su parte principal, es decir, por la ecuación

$$ax^2 + 2bxy + cy^2 = 0,$$

que para el caso dado es precisamente la ecuación (41.106), pues aquí se ve claramente que

$$a = F''_{xx}, \quad b = F''_{xy} \quad \text{y} \quad c = F''_{yy}.$$

En cambio, si el punto (0, 0) satisface la ecuación (41.107), pero no es un punto singular, entonces (41.107) tiene por expresión

$$Ax + By + R(x, y) = 0, \quad A^2 + B^2 > 0,$$

donde $R(x, y)$ es un polinomio, todos los términos del cual tienen el orden no inferior al segundo. Del teorema sobre las funciones implícitas (véase el teorema 1 en el p. 41.1) se deduce que la ecuación

$$Ax + By = 0$$

es en este caso una ecuación de la tangente en el punto (0, 0) a la gráfica de la solución de la ecuación (41.107).

Ejercicios. Investíguese el comportamiento de cada una de las siguientes curvas en el entorno de sus puntos singulares; hállense las tangentes en el punto singular.

13. $y^2 = x^2 + x^3$. 14. $y^2 = x^2 - x^3$.

15. $y^2 = x^2 - x^4$.
16. $(x^2 - 9)y^2 = x^4$.
17. $y^2 = x(x - 3)^2$.
18. $y(y - 2)^2 = x^2$.
19. $4y^2 = x^5 + 5x^4$.
20. $(x^2 - y^2)y = x^4$.
21. $(y - x^2)^2 = x^5$.

41.10. CAMBIO DE VARIABLES

En las diversas cuestiones del análisis matemático y en las aplicaciones de éste, al estudiar tal o cual fórmula que contiene unas funciones y sus derivadas (ordinarias o parciales), resulta a menudo conveniente pasar a otras variables independientes y, a veces, a otras funciones ligadas por relaciones determinadas con las funciones que figuran en la fórmula que se considera. Todas estas transformaciones se realizan sobre la base de las reglas de derivación de las funciones implícitas y compuestas. He aquí unos ejemplos.

Sea $u = u(x, y)$. Transformemos las expresiones

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \text{ y } \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2},$$

reduciéndolas a las coordenadas polares r y φ. La primera de estas expresiones es el cuadrado de la longitud del gradiente ∇u de la función u, es decir, es igual a $|\nabla u|^2$, mientras que la segunda expresión tiene una designación especial Δu:

$$|\nabla u|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \tag{41.108}$$

$$\Delta u \stackrel{\text{def}}{=} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{41.109}$$

El símbolo Δ, que es indicio de la aplicación a la función u de la operación (41.109), lleva el nombre de *operador de Laplace**).

De las fórmulas que ligan las coordenadas cartesianas con las polares

$$x = r\cos\varphi, y = r\operatorname{sen}\varphi, \tag{41.110}$$

encontramos

$$\frac{\partial x}{\partial r} = \cos\varphi, \frac{\partial x}{\partial \varphi} = -r\operatorname{sen}\varphi, \frac{\partial y}{\partial r} = \operatorname{sen}\varphi, \frac{\partial y}{\partial \varphi} = r\cos\varphi. \tag{41.111}$$

Apliquemos las fórmulas de derivación de una función compuesta:

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial u}{\partial x}\cos\varphi + \frac{\partial u}{\partial y}\operatorname{sen}\varphi,$$

$$\frac{\partial u}{\partial \varphi} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \varphi} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \varphi} = -\frac{\partial u}{\partial x}r\operatorname{sen}\varphi + \frac{\partial u}{\partial y}r\cos\varphi.$$

*) P. Laplace (1749 — 1827), mecánico y matemático francés.

Resolvamos estas igualdades respecto de $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$:

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}, \quad \frac{\partial u}{\partial y} = \frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r} \tag{41.112}$$

y sustituyamos las expresiones obtenidas en (41.108):

$$|\nabla u|^2 = \left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)^2 + \left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)^2 =$$

$$= \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \varphi}\right)^2.$$

Pasemos ahora al cálculo de la expresión (41.109). Derivemos las fórmulas (41.110) respecto de x y luego respecto de y:

$$\left.\begin{aligned} 1 &= \cos\varphi\frac{\partial r}{\partial x} - r\operatorname{sen}\varphi\frac{\partial \varphi}{\partial x}, \\ 0 &= \operatorname{sen}\varphi\frac{\partial r}{\partial x} + r\cos\varphi\frac{\partial \varphi}{\partial x}, \end{aligned}\right\} \quad \left.\begin{aligned} 0 &= \cos\varphi\frac{\partial r}{\partial y} - r\operatorname{sen}\varphi\frac{\partial \varphi}{\partial y}, \\ 1 &= \operatorname{sen}\varphi\frac{\partial r}{\partial y} + r\cos\varphi\frac{\partial \varphi}{\partial y}, \end{aligned}\right\}$$

Resolvamos los sistemas obtenidos respecto de $\frac{\partial r}{\partial x}$, $\frac{\partial \varphi}{\partial x}$, $\frac{\partial r}{\partial y}$ y $\frac{\partial \varphi}{\partial y}$;

$$\frac{\partial r}{\partial x} = \cos\varphi, \frac{\partial r}{\partial y} = \operatorname{sen}\varphi, \frac{\partial \varphi}{\partial x} = -\frac{\operatorname{sen}\varphi}{r}, \frac{\partial \varphi}{\partial y} = \frac{\cos\varphi}{r}. \tag{41.113}$$

Derivemos ahora las fórmulas (41.112) respecto de x e y; entonces, empleando (41.113), obtendremos

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)\frac{\partial r}{\partial x} + \frac{\partial}{\partial \varphi}\left(\frac{\partial u}{\partial r}\cos\varphi - \right.$$

$$\left. - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)\frac{\partial \varphi}{\partial x} = \frac{\partial^2 u}{\partial r^2}\cos^2\varphi - \frac{2\cos\varphi\operatorname{sen}\varphi}{r}\frac{\partial^2 u}{\partial r\partial \varphi} +$$

$$+ \frac{\operatorname{sen}^2\varphi}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\operatorname{sen}^2\varphi}{r}\frac{\partial u}{\partial r} + \frac{2\cos\varphi\operatorname{sen}\varphi}{r^2}\frac{\partial u}{\partial \varphi};$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)\frac{\partial r}{\partial y} +$$

$$+ \frac{\partial}{\partial \varphi}\left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)\frac{\partial \varphi}{\partial y} = \frac{\partial^2 u}{\partial r^2}\operatorname{sen}^2\varphi +$$

$$+ \frac{2\cos\varphi\operatorname{sen}\varphi}{r}\frac{\partial^2 u}{\partial r\partial \varphi} + \frac{\cos^2\varphi}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\cos^2\varphi}{r}\frac{\partial u}{\partial r} - \frac{2\cos\varphi\operatorname{sen}\varphi}{r^2}\frac{\partial u}{\partial \varphi}.$$

Al sustituir las expresiones obtenidas en (41.109), tendremos

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{1}{r}\frac{\partial u}{\partial r}.$$

Cuando en la expresión a transformar figuran no una sola, sino varias derivadas de orden dado, resulta más cómodo emplear el método de cálculo de las diferenciales, en lugar de las derivadas. Por ejemplo, considerando x e y como variables independientes, obtendremos las expresiones para las diferenciales dr y $d\varphi$. De las fórmulas (41.110) tenemos

$$dx = \cos\varphi\, dr - r \operatorname{sen}\varphi\, d\varphi,\ dy = \operatorname{sen}\varphi\, dr + r\cos\varphi\, d\varphi,$$

de aquí

$$dr = \cos\varphi dx + \operatorname{sen}\varphi\, dy,\ d\varphi = -\frac{\operatorname{sen}\varphi}{r}dx + \frac{\cos\varphi}{r}dy \qquad (41.114)$$

(observemos que de estas fórmulas se deducen inmediatamente también las fórmulas (41.113).

Para la función $u = u(x, y)$ tenemos

$$du = \frac{\partial u}{\partial r}dr + \frac{\partial u}{\partial \varphi}d\varphi =$$

$$= \left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)dx + \left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)dy. \quad (41.115)$$

En la expresión para la diferencial du los coeficientes de dx y dy son las derivadas $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$, por lo cual, de (41.115) se obtienen las dos fórmulas (41.112).

Hallemos, a continuación, las segundas diferenciales d^2r y $d^2\varphi$, haciendo uso de (41.114):

$$d^2r = -\operatorname{sen}\varphi\, d\varphi\, dx + \cos\varphi\, d\varphi\, dy =$$

$$= \frac{\operatorname{sen}^2\varphi\, dx^2 - 2\cos\varphi \operatorname{sen}\varphi\, dx\, dy + \cos^2\varphi\, dy^2}{r},$$

$$d^2\varphi = -\left(\frac{\cos\varphi}{r}dx + \frac{\operatorname{sen}\varphi}{r}dy\right)d\varphi + \left(\frac{\operatorname{sen}\varphi}{r^2}dx - \frac{\cos\varphi}{r^2}dy\right)dr =$$

$$= \frac{2\cos\varphi \operatorname{sen}\varphi\, dx^2 - 2(\cos^2\varphi - \operatorname{sen}^2\varphi)\, dx\, dy - 2\cos\varphi \operatorname{sen}\varphi\, dy^2}{r^2}.$$

Ahora, de (41.115) obtendremos para d^2u

$$d^2u = \frac{\partial^2 u}{\partial r^2}dr^2 + 2\frac{\partial^2 u}{\partial r\, d\varphi}dr\, d\varphi + \frac{\partial^2 u}{\partial \varphi^2}d\varphi^2 + \frac{\partial u}{\partial r}d^2r + \frac{\partial u}{\partial \varphi}d^2\varphi =$$

$$= \left(\cos^2\varphi\,\frac{\partial^2 u}{\partial r^2} - \frac{2\cos\varphi\,\operatorname{sen}\varphi}{r}\,\frac{\partial^2 u}{\partial r\partial\varphi} + \frac{\operatorname{sen}^2\varphi}{r^2}\,\frac{\partial^2 u}{\partial\varphi^2} + \right.$$

$$\left. + \frac{\operatorname{sen}^2\varphi}{r}\,\frac{\partial u}{\partial r} + \frac{2\cos\varphi\operatorname{sen}\varphi}{r^2}\,\frac{\partial u}{\partial\varphi}\right) dx^2 + 2(\ldots)\,dx\,dy + (\ldots)\,dy^2.$$

De aquí se obtienen las expresiones para las segundas derivadas $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x\partial y}$ y $\frac{\partial^2 u}{\partial y^2}$ que intervienen como los coeficientes de dx^2, $2dx\,dy$ y dy^2.

Los métodos análogos son aplicables, por supuesto, en el caso en que se realiza algún otro cambio de variables $x = x(u, v)$, $y = y(u, v)$, cuando se tienen derivadas de órdenes superiores y también cuando se trata de las funciones del mayor número de variables.

Ejercicio 22. Transfórmese la expresión $|\nabla u|^2$, donde $u = u(x, y)$ a las coordenadas ortogonales ξ, η, es decir, tales que

$$\frac{\partial x}{\partial\xi}\,\frac{\partial x}{\partial\eta} + \frac{\partial y}{\partial\xi}\,\frac{\partial y}{\partial\eta} = 0.$$

23. Transfórmese la ecuación $y'' - xy'^3 + e^y y'^3 = 0$, considerando y como una nueva variable independiente y x, como una función de y.

24. En la ecuación $\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = 0$ pásese a las nuevas variables independientes $u = x + y$, $v = x - y$.

25. En la expresión $-\frac{1}{2}z\left(\frac{1}{x^2}\,\frac{\partial^2 z}{\partial x^2} + \frac{1}{y^2}\,\frac{\partial^2 z}{\partial y^2}\right) + \frac{1}{2}\left[\left(\frac{1}{x}\,\frac{\partial z}{\partial x}\right)^2 + \left(\frac{1}{y}\,\frac{\partial z}{\partial y}\right)^2\right] -$

$- \frac{1}{2}z\left(\frac{1}{x^3}\,\frac{\partial z}{\partial x} + \frac{1}{y^3}\,\frac{\partial z}{\partial y}\right)$ pásese a las variables u, v, $w = w(u, v)$, si $u = x^2$, $v = y^2$, $w = z^2$ (respuesta: $\frac{\partial^2 w}{\partial u^2} + \frac{\partial^2 w}{\partial v^2}$).

Problema 27. En un espacio n-dimensional transfórmese la expresión $|\nabla u|^2$, donde $u = u(x_1, \ldots, x_n)$, reduciéndola a las coordenadas ortogonales $\xi_1, \ldots, \xi_n$, es decir, a tales coordenadas que para $i \neq k$ se verifica la igualdad

$$\sum_{j=1}^{n} \frac{\partial x_j}{\partial\xi_i}\,\frac{\partial x_j}{\partial\xi_k} = 0, \quad i, k = 1, 2, \ldots, n.$$

§ 42. DEPENDENCIA DE LAS FUNCIONES

42.1. CONCEPTO DE DEPENDENCIA DE LAS FUNCIONES. CONDICIÓN NECESARIA PARA LA DEPENDENCIA DE LAS FUNCIONES

Definición 1. *Sean dadas en el conjunto abierto $G \subset R^n$ las funciones continuamente derivables*

$$y_i = \varphi_i(x), i = 1, 2, \ldots, m, x = (x_1, \ldots, x_n) \in G. \quad (42.1)$$

Si existen un conjunto abierto D en el espacio $R^{m-1}_{y_1, \ldots, y_{m-1}}$ y una función, continuamente derivable en D, $\Phi(y_1, \ldots, y_{m-1})$ tales que en todo punto $x \in G$ se cumplen las condiciones $(\varphi_1(x), \ldots, \varphi_{m-1}(x)) \in D$ y $\Phi(\varphi_1(x), \ldots, \varphi_{m-1}(x)) = \varphi_m(x)$, entonces la función φ_m se llama dependiente en el conjunto G de las funciones $\varphi_1, \ldots, \varphi_{m-1}$.

Definición 2. *Si entre las funciones del sistema* (42.1) *hay una función, dependiente de las demás en el conjunto G, dicho sistema se denomina dependiente en el conjunto G.*

Si ninguna función del sistema (42.1) *depende de las restantes en el conjunto G, este sistema se llama independiente en G.*

A veces, para abreviar, en lugar de la expresión "sistema de funciones dependiente (independiente)" diremos simplemente "funciones dependientes (independientes)".

En la cuestión sobre la dependencia del sistema de funciones (42.1) el papel fundamental lo desempeña la matriz de Jacobi de este sistema

$$\left\| \frac{\partial y_i}{\partial x_j} \right\|, i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n, \quad (42.2)$$

i es el número de la fila y j, el número de la columna.

Teorema 1 (condición necesaria para la dependencia de las funciones). *Supongamos que $m \leqslant n$ y el sistema de funciones* (42.1) *es dependiente en un conjunto abierto G. Entonces, en todo punto de este conjunto el rango de la matriz de Jacobi* (42.2)*) *del sistema citado es inferior a m.*

DEMOSTRACIÓN. Por hipótesis, el sistema de funciones (42.1) es dependiente en G, puesto que por lo menos una de estas funciones depende de las demás. Sea, para concretar, que φ_m depende de $\varphi_1, \ldots, \varphi_{m-1}$:

$$\varphi_m(x) = \Phi(\varphi_1(x), \ldots, \varphi_{m-1}(x)), x \in G,$$

donde Φ es una función continuamente derivable de $(m-1)$ argumentos $y_1, \ldots, y_{m-1}$. De aquí

$$\frac{\partial y_m}{\partial x_j} = \sum_{i=1}^{m-1} \frac{\partial \Phi}{\partial y_i} \frac{\partial y_i}{\partial x_j} \text{ para cualquier } j = 1, 2, \ldots, n.$$

Esta fórmula muestra que la m-ésima fila de la matriz de Jacobi (42.2) en todo punto $x \in G$ es una combinación lineal de las filas restantes de esta matriz y, por lo tanto, el rango de la matriz de Jacobi (42.2) es inferior a m en todo punto $x \in G$. □

Corolario. *Supongamos que $m = n$ y el sistema de funciones* (42.1) *es depen-*

*) Recordemos que se llama rango de una matriz el número máximo de sus filas linealmente independientes. Este número coincide con el orden máximo del menor de dicha matriz distinto de cero.

diente en G. En este caso su jacobiano $\dfrac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}$ *es igual a cero en todos los puntos del conjunto G.*

Gorolario 2 (condiciones suficientes para la independencia de las funciones). *Supongamos que* $m \leqslant n$ *y el rango de la matriz de Jacobi* (42.2) *por lo menos en un punto del conjunto abierto G es igual a m. En este caso el sistema* (42.1) *es independiente en el conjunto G.*

El corolario 1 se obtiene inmediatamente del teorema demostrado cuando $m = n$.

El corolario 2 se demuestra con facilidad por reducción al absurdo.

Por cuanto las filas de la matriz de Jacobi (42.2) son coordenadas de los gradientes de la función (42.1), el teorema 1 puede parafrasearse del modo siguiente.

Si el sistema de funciones (42.1) *es dependiente en la región G, los gradientes* $\nabla\varphi_1, \ldots, \nabla\varphi_m$ *de estas funciones son linealmente dependientes en todo punto de G.*

42.2. CONDICIONES SUFICIENTES PARA LA DEPENDENCIA DE LAS FUNCIONES

En este punto conservamos intactas las designaciones del párrafo anterior y supondremos, como hasta ahora, que las funciones (42.1) son continuamente derivables en el conjunto abierto $G \subset R^n$.

Teorema 2 (condiciones suficientes para la dependencia local de las funciones). *Supongamos que el rango de la matriz de Jacobi* (42.2) *del sistema de funciones* (42.1) *en cualquier punto del conjunto abierto G no es superior al número r,* $r < m \leqslant n$, *y en cierto punto* $x^{(0)} \in G$ *es igual a r, en otras palabras, existen tales variables* $x_{j_1}, \ldots, x_{j_r}$ *y las funciones* $y_{i_1} = \varphi_{i_1}(x), \ldots, y_{i_r} = \varphi_{i_r}(x)$ *que*

$$\left.\frac{\partial(y_{i_1}, \ldots, y_{i_r})}{\partial(x_{j_1}, \ldots, x_{j_r})}\right|_{x^{(0)}} \neq 0. \tag{42.3}$$

En este caso todas r funciones, que figuran en la condición (42.3), *son independientes en el conjunto G y existe un entorno del punto* $x^{(0)}$ *tal que cualquiera de* $m - r$ *funciones restantes depende en dicho entorno de las r funciones mencionadas.*

DEMOSTRACIÓN. Supongamos, para simplificar la anotación, que la condición (42.3) tiene por expresión

$$\left.\frac{\partial(y_1, \ldots, y_r)}{\partial(x_1, \ldots, x_r)}\right|_{x^{(0)}} \neq 0 \tag{42.4}$$

(esto siempre puede lograrse, numerando, si se desea, las funciones y los argumentos del sistema (42.1) en el orden debido). De acuerdo con el corolario 2 del teorema 1 en el p. 42.1, las funciones $y_1, \ldots, y_r$ son independientes en G.

Mostremos que cada una de las funciones restantes depende de ellas en cierto entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$. Sea $y_i^{(0)} = \varphi_i(x^{(0)})$, $i = 1, 2, \ldots, m$. Examinemos el sistema de las primeras r funciones del sistema (42.1):

$$\begin{aligned} y_1 &= \varphi_1(x_1, \ldots, x_n) \\ &\cdots\cdots\cdots\cdots \\ y_r &= \varphi_r(x_1, \ldots, y_n)\,. \end{aligned} \tag{42.5}$$

Elijamos ante todo tal η_0 que todo punto $x = (x_1, \dots, x_n)$ perteneciente a η_0-entorno cúbico del punto $x^{(0)}$, es decir, todo punto x, para el cual $|x_i - x_i^{(0)}| < \eta_0$, $i = 1, 2, \dots, n$, pertenezca al conjunto G: $x \in G$. Esto es siempre posible, pues el conjunto mencionado es abierto.

Luego, en virtud de la condición (42.4) y el teorema sobre las funciones implícitas (véase el p. 41.3), el sistema (42.5) es resoluble respecto de las variables $x_1, \dots, x_r$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$:

$$\begin{aligned} x_1 &= f_1(y_1, \dots, y_r, x_{r+1}, \dots, x_n), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ x_r &= f_r(y_1, \dots, y_r, x_{r+1}, \dots, x_n). \end{aligned} \tag{42.6}$$

Las funciones $f_1, \dots, f_r$ en este caso están definidas y son continuamente derivables en cierto entorno del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)})$.

En la forma más amplia (si a título de entorno se toman los entornos cúbicos) esto significa lo siguiente: pueden escogerse tales números $\delta > 0$ y $\eta > 0$ (con el fin de comodidad, pueden ser incluso inferiores a η_0: $\delta < \eta_0$, $\eta < \eta_0$) que si designamos con U el entorno cúbico del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)})$ prefijado por las desigualdades

$$|y_i - y_i^{(0)}| < \delta, \quad i = 1, 2, \dots, r, \quad |x_j - x_j^{(0)}| < \delta, \quad j = r+1, \dots, n$$

entonces

1) en el entorno U las funciones f_k, $k = 1, 2, \dots, r$, están definidas y son continuamente derivables;

2) para todos los puntos $(y_1, \dots, y_r, x_{r+1}, \dots, x_n) \in U$ son válidas las desigualdades

$$|f_k(y_1, \dots, y_r, x_{r+1}, \dots, x_n) - x_k^{(0)}| < \eta, \quad k = 1, 2, \dots, r;$$

3) en el entorno U se verifican las igualdades

$$\varphi_i(f_1, \dots, f_r, x_{r+1}, \dots, x_n) = y_i, \quad i = 1, 2, \dots, r,$$

donde por f_k, $k = 1, \dots, r$, se entienden los segundos miembros de las igualdades (42.6).

Consideremos una composición de funciones (42.6) y $\varphi_{r+1}(x_1, \dots, x_n)$, es decir, una función

$$y_{r+1} = \varphi_{r+1}(f_1, \dots, f_r, x_{r+1}, \dots, x_n), \tag{42.7}$$

en la que $f_k = f_k(y_1, \dots, y_r, x_{r+1}, \dots, x_n)$, $k = 1, \dots, r$. Esta función compuesta está definida a ciencia cierta y es continuamente derivable en el entorno cúbico U, mencionado arriba, del punto

$$(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)}).$$

Probemos que, realmente, la función (42.7) en este entorno U no depende de las variables $x_{r+1}, \dots, x_n$, es decir, no varía cuando cambian las variables citadas y, consecuentemente, es, de hecho, sólo función de las variables $y_1, \dots, y_r$. Con este objeto basta mostrar que para la función (42.7) en el entorno U se verifica la igualdad

$$\frac{\partial y_{r+1}}{\partial x_j} = 0, j = r+1, \dots, n \tag{42.8}$$

(véase el p. 20.4 o la fórmula de los incrementos finitos de Lagrange en el p. 39.2, de la cual se deduce directamente la suficiencia de la condición (42.8) para que una función sea independiente de las variables $x_{r+1}, \ldots, x_n$ en una región convexa y, por consiguiente, en un entorno cúbico).

Para demostrar la igualdad (42.8), fijemos uno de los valores $j (j = r+1, \ldots, n)$ y las coordenadas x_k cuyos índices k toman los valores $r+1, \ldots, j-1, j+1, \ldots, n$; designémoslas mediante x_k^* y elijamos x_k^* de una manera tal que sea $|x_k^* - x_k^{(0)}| < \delta$, $k = r+1, \ldots, j-1, j+1, \ldots, n$.

Consideraremos una aplicación

$$\begin{gathered} y_1 = y_1, \\ \ldots\ldots\ldots \\ y_r = y_r, \\ y_{r+1} = \varphi_{r+1}(f_1^*, \ldots, f_r^*, x_{r+1}^*, \ldots, x_{j-1}^*, x_j, x_{j+1}^*, \ldots, x_n^*), \end{gathered} \tag{42.9}$$

donde $f_k^* = f_k(y_1, \ldots, y_r, x_{r+1}^*, \ldots, x_{j-1}^*, x_j, x_{j+1}^*, \ldots, x_n^*)$, del entorno cúbico $U^{(0)}$ del punto $(y_1^{(0)}, \ldots, y_r^{(0)}, x_j^{(0)})$ que se define por las desigualdades

$$|y_k - y_k^{(0)}| < \delta, \; k = 1, 2, \ldots, r, \; |x_j - x_j^{(0)}| < \delta.$$

Para subrayar cuáles son precisamente las variables que cambian, representemos simbólicamente la aplicación (42.9) en la forma

$$(y_1, \ldots, y_r, x_j) \to (y_1, \ldots, y_r, y_{r+1}).$$

Esta aplicación es continuamente derivable en $U^{(0)}$; su matriz de Jacobi tiene por expresión

$$\left\| \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \dfrac{\partial y_{r+1}}{\partial y_1} & \dfrac{\partial y_{r+1}}{\partial y_2} & \dfrac{\partial y_{r+1}}{\partial y_r} & \dfrac{\partial y_{r+1}}{\partial x_j} \end{matrix} \right\|$$

y, por esta razón

$$\frac{\partial(y_1, \ldots, y_r, y_{r+1})}{\partial(y_1, \ldots, y_r, x_j)} = \frac{\partial y_{r+1}}{\partial x_j}, \tag{42.10}$$

es decir, el jacobiano de la aplicación en consideración es igual a la derivada que nos interesa.

En el entorno $U^{(0)}$ esta aplicación puede representarse como composición de dos aplicaciones: de la aplicación continuamente derivable

$$\begin{gathered} x_1 = f_1(y_1, \ldots, y_r, x_{r+1}^*, \ldots, x_{j-1}^*, x_j, x_{j+1}^*, \ldots, x_n^*), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = f_r(y_1, \ldots, y_r, x_{r+1}^*, \ldots, x_{j-1}^*, x_j, x_{j+1}^*, \ldots, x_n^*), \\ x_j = x_j \end{gathered}$$

del entorno $U^{(0)}$ y la aplicación continuamente derivable

$$y_1 = \varphi_1(x_1, \dots, x_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$y_r = \varphi_r(x_1, \dots, x_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*),$$

$$y_{r+1} = \varphi_{r+1}(x_1, \dots, x_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*)$$

del entorno del punto $(x_1^{(0)}, \dots, x_r^{(0)}, x_j^{(0)})$ definido mediante las desigualdades

$$|x_i - x_i^{(0)}| < \eta,\ i = 1, 2, \dots, r,\ |x_j - x_j^{(0)}| < \delta.$$

Por ser adecuadamente elegidos los números δ y η, la composición de estas aplicaciones la que con los fines ilustrativos puede ser representada en la forma

$$(y_1, \dots, y_r, x_j) \to (x_1, \dots, x_r, x_j) \to (y_1, \dots, y_r, y_{r+1}),$$

está definida y es continuamente derivable en el entorno $U^{(0)}$. La primera de estas aplicaciones es continuamente derivable en el entorno $U^{(0)}$ del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_j^{(0)})$ y la segunda es continuamente derivable en el entorno correspondiente del punto $(x_1^{(0)}, \dots, x_r^{(0)}, x_j^{(0)})$. Por ello, de (42.10) y de las propiedades de los jacobianos de aplicaciones (véase el p. 41.7) tenemos

$$\frac{\partial y_{r+1}}{\partial x_j} = \frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(y_1, \dots, y_r, x_j)} = \frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(x_1, \dots, x_r, x_j)} \frac{\partial(x_1, \dots, x_r, x_j)}{\partial(y_1, \dots, y_r, x_j)}. \quad (42.11)$$

De acuerdo con la condición del teorema, el rango de la matriz de Jacobi en el conjunto G es igual o inferior a r, por consiguiente,

$$\frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(x_1, \dots, x_r, x_j)} = 0$$

en todo punto de G. Por ello, de (42.11) se desprende en seguida que para cualquier punto $(y_1, \dots, y_r, x_j) \in U^{(0)}$ y, por ende, para todo punto

$$(y_1, \dots, y_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*) \in U$$

se verifica la igualdad (42.8). Puesto que las coordenadas x_k^* se han fijado arbitrariamente, con tal de que sea $|x_k^* - x_k^{(0)}| < \delta$, $k = r+1, \dots, j-1, j+1, \dots, n$, esto quiere decir que la igualdad (42.8) es justa en todo el entorno U.

De este modo, la función (42.7) depende sólo de las variables $y_1, \dots, y_r$. Al designarla por el símbolo Φ, obtendremos

$$\varphi_{r+1}(f_1, \dots, f_r, x_{r+1}, \dots, x_n) = \Phi(y_1, \dots, y_r).$$

Elijamos ahora δ_0, $\delta_0 < \delta$ y $\delta_0 < \eta$, de una manera tal que para $|x_i - x_i^{(0)}| < \delta_0$, $i = 1, 2, \dots, n$, se cumplan las desigualdades

$$|y_i - y_i^{(0)}| < \delta,\ j = 1, 2, \dots, r.$$

Esto es posible, puesto que las funciones $y_j = \varphi_j(x_1, \dots, x_n)$, $j = 1, 2, \dots, r$, del sistema (42.5) son continuas en el punto $x^{(0)}$.

En virtud de lo demostrado, para cualquier punto x del entorno δ_0-cúbico del punto $x^{(0)}$, es decir, para todo punto $x = (x_1, \ldots, x_n)$ tal que

$$|x_i - x_i^{(0)}| < \delta_0, \quad i = 1, 2, \ldots, n,$$

será válida la identidad

$$\varphi_{r+1}(x) = \Phi(\varphi_1(x), \ldots, \varphi_r(x)),$$

es decir, en el entorno citado del punto $x^{(0)}$ las funciones $\varphi_1, \ldots, \varphi_r, \varphi_{r+1}$ son dependientes.

Análogamente se demuestra también la dependencia de cada una de las funciones $\varphi_{r+2}, \ldots, \varphi_m$ de $\varphi_1, \ldots, \varphi_r$ en cierto entorno del punto $x^{(0)}$. □

Al igual que la condición necesaria para la dependencia de las funciones, las condiciones suficientes pueden enunciarse también en términos de los gradientes. Para simplificar, limitémonos al caso en que $r = m - 1$.

Si los gradientes $\nabla\varphi_1, \ldots, \nabla\varphi_m$ son linealmente dependientes en todo punto de la región G, entonces, cualquiera que sea el punto $x \in G$ en el que $m - 1$ de los gradientes mencionados son linealmente independientes, existe un entorno suyo en el cual las funciones $\varphi_1, \ldots, \varphi_m$ son dependientes. Además, si, por ejemplo, los gradientes $\nabla\varphi_1, \ldots, \nabla\varphi_{m-1}$ son linealmente independientes en el punto dado y, por lo tanto, el gradiente $\nabla\varphi_m$ en él es una combinación lineal de ellos, entonces en el entorno mencionado la función φ_m depende de las funciones $\varphi_1, \ldots, \varphi_{m-1}$.

Conviene fijar la atención en que las condiciones suficientes para la dependencia de las funciones, establecidas aquí, tienen un carácter local, a diferencia de los resultados del punto anterior que son de carácter global. Esto significa lo siguiente: si un sistema de m funciones continuamente derivables (42.1) es dependiente en el conjunto abierto $G \subset R^n$, entonces, de acuerdo con el teorema 1 del p. 42.1, en *todo* punto de dicho conjunto el rango de la matriz de Jacobi de este sistema es inferior a m (correspondientemente, si por lo menos en un punto del cunjunto G el rango de la matriz que se considera es igual a m, el sistema es independiente en *todo* el conjunto G). En cuanto al teorema 2 de este punto, se debe decir que éste sólo afirma que si en algún punto $x^{(0)} \in G$ se cumplen sus condiciones, el sistema dado de funciones será dependiente sólo en *cierto entorno* del punto citado (y no en todo el conjunto G). De este modo, efectivamente, la afirmación del teorema 2 tiene un carácter local.

Diremos, además, que si en *todo* punto $x^{(0)}$ del conjunto abierto G se cumplen las condiciones del teorema 2, entonces, naturalmente, en este caso, el sistema de funciones en consideración será dependiente en cierto entorno de todo punto. No obstante, el teorema 2 no garantiza que dicha dependencia será la misma en todos los entornos indicados, es decir, del teorema 2 no se infiere que en diferentes puntos unas mismas funciones serán dependientes de las otras y que las funciones Φ que "cristalizan en realidad" la dependencia de unas mismas funciones, consideradas en distintos entornos, serán coincidentes en los puntos de intersección de estos entornos. Por consiguiente, del teorema 2 no se infiere ni mucho menos que el sistema de funciones que satisface las condiciones de este teorema en todos los puntos $x^{(0)}$ del conjunto G será dependiente en todo el conjunto G en su total en el único sentido, es decir, en el sentido de la definición 1. Esto precisamente sirve de testimonio de que el teorema 2 no lleva el carácter global.

Observemos que existe un acceso algo más general al concepto de dependencia de las funciones que permite construir la teoría global de este problema, sin embargo no nos detendremos en ello.

Ejemplo. Examinemos un sistema de funciones

$$u = \operatorname{sen}(x + y)$$
$$v = \cos(x + y), \tag{42.12}$$

El jacobiano de este sistema es nulo en todo el plano

$$\begin{vmatrix} \cos(x+y) & \cos(x+y) \\ -\operatorname{sen}(x+y) & -\operatorname{sen}(x+y) \end{vmatrix} = 0,$$

y, como es fácil de ver, el rango de la matriz de Jacobi de este sistema es igual a uno en todos los puntos del plano.

Según el teorema 2, las funciones (42.12) son dependientes en el entorno de cada punto del plano. En el caso dado la dependencia de las funciones se halla fácilmente en la forma explícita, por ejemplo, en el conjunto abierto de los puntos (x, y), para los cuales $\cos(x + y) > 0$, ella puede ser dada por la fórmula $v = \sqrt{1 - u^2}$.

Ejercicios. 1. Sean $u = x^2 + y^2 + z^2$, $v = xy + yz + zx$. $w = x + y + z$. Demuéstrese que las funciones u, v, w son dependientes y hállese la ecuación que expresa su dependencia.

2. Investíguese la cuestión sobre la dependencia de las funciones $u = \xi^3 + \eta^3 + \zeta^3$, $v = \xi\eta\zeta$, $w = \xi^2 + \eta^2 + \zeta^2$, $z = \xi\eta + \eta\zeta + \zeta\xi$.

Problema 28. Una función $u = u(x, y)$ se denomina armónica en una región plana, si en todos los puntos de esta región satisface la ecuación $\Delta u = 0$ (véase (41.109)). Demuéstrese que dos funciones armónicas son dependientes en una región plana cuando, y sólo cuando son linealmente dependientes.

§ 43. EXTREMO CONDICIONADO

43.1. CONCEPTO DE EXTREMO CONDICIONADO

Supongamos que en un conjunto abierto $G \subset R^n$ vienen definidas las funciones

$$y_i = f_i(x), i = 1, 2, \ldots, m, \tag{43.1}$$

$x = (x_1, \ldots, x_n) \in G$. Designemos mediante X un conjunto de los puntos $x \in G$, donde todas las funciones f_i, $i = 1, 2, \ldots, m$ se reducen a cero:

$$X = \{x : f_i(x) = 0, i = 1, 2, \ldots, m, x \in G\}. \tag{43.2}$$

Las ecuaciones

$$f_i(x) = 0, i = 1, 2, \ldots, m, \tag{43.3}$$

se llamarán *ecuaciones de conexión (enlace)*.

Definición 1. *Sea $y = f_0(x)$ una función definida en G. El punto $x^{(0)} \in X$ se llama punto de extremo condicionado*) de la funcion $f_0(x)$ respecto de las ecuaciones de enlace (43.3) (o si se cumplen dichas ecuaciones), siempre que es un punto de extremo ordinario de esta función al considerarse ésta sólo en el conjunto X* (véase el p. 40.1).

En otras palabras, el valor de la función $f_0(x)$ en el punto $x^{(0)}$ se compara aquí no con todos los valores suyos en un entorno suficientemente pequeño de este punto, sino sólo con los valores en los puntos pertenecientes, a la vez, al mencionado entorno suficiente pequeño y al conjunto X. Al igual que en el caso de los extremos ordinarios, podemos, desde luego, considerar los puntos de extremo simplemente condicionado y los de extremo estrictamente condicionado.

Ejemplos. 1. Consideraremos una función

$$f(x, y) = x^2 + y^2 \tag{43.4}$$

y la ecuación de enlace

$$x + y - 1 = 0. \tag{43.5}$$

Hallemos el extremo condicionado de la función (43.4) para el caso en que se cumple la ecuación de enlace (43.5). De (43.5) tenemos $y = 1 - x$, de donde

$$f(x, 1 - x) = 2x^2 - 2x + 1.$$

De este modo, cuando se cumple la condición de enlace, la función (43.4) es una función de una sola variable. Su extremo se halla de una manera elemental: igualando a cero su derivada (la condición necesaria de un extremo), obtenemos $2x - 1 = 0$, de donde $x = \frac{1}{2}$. En el punto dado la función que se considera, tiene, evidentemente, un mínimo (la función es un polinomio de segundo grado cuyo término mayor tiene coeficiente positivo). Al valor $x = \frac{1}{2}$, de conformidad con la ecuación de enlace (43.5), le corresponde $y = \frac{1}{2}$.

Por consiguiente, en el punto (1/2, 1/2) la función (43.4) alcanza su mínimo respecto de la ecuación de enlace (43.5). En el lenguaje geométrico esto significa que el punto del paraboloide $z = x^2 + y^2$, que se proyecta en el punto (1/2, 1/2), es el más inferior de todos sus puntos dispuestos por encima de la recta (43.5) (fig. 169). Este ejemplo muestra que un punto en el que la función alcanza su extremo condicionado, no es, en el caso general, punto de extremo de esta función.

2. Consideraremos la función $f(x, y) = y^2 - x^2$ y la ecuación de enlace $y = 2x$.

Tenemos $f(x, 2x) = 3x^2$, es decir, si se cumplen las ecuaciones de enlace, la función en consideración es también una función de una sola variable y, evidentemente, alcanza su mínimo cuando $x = 0$ (fig. 170). De conformidad con la ecuación de enlace, al valor $x = 0$ corresponde el valor de $y = 0$, y, por lo tanto, la función $f(x, y) = y^2 - x^2$ tiene en el punto (0, 0) un mínimo condicionado respecto de la ecuación de enlace $y = 2x$.

*) Se ha adoptado también el término "extremo relativo".

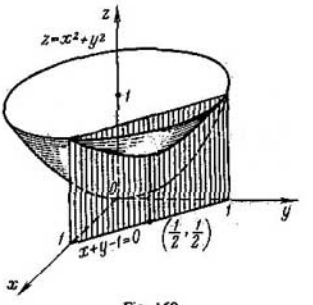


Fig. 169

Cabe destacar que en este caso la propia función $f(x, y)$ no tiene máximo ni mínimo en ningún punto del plano. De este modo, el ejemplo considerado muestra que la función puede no tener un extremo, pero, para ciertas ecuaciones de enlace, puede tener un extremo condicionado.

En lo sucesivo supondremos que

1) todas las funciones $f_0, f_1, \ldots, f_m$ son continuamente derivables en el conjunto abierto G;

2) en el punto $x^{(0)}$ que se considera, los vectores $\nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, es decir, el rango de la matriz de Jacobi

$$\left(\frac{\partial f_j}{\partial x_i}\right), j = 1, 2, \ldots, m, i = 1, 2, \ldots, n,$$

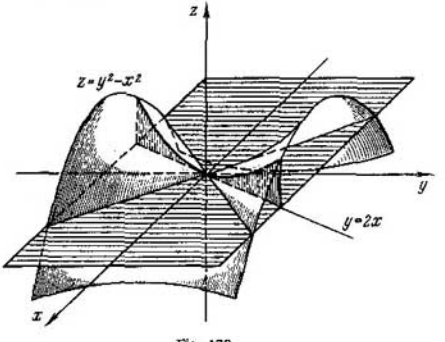


Fig. 170

es igual a m, o sea, al número de sus filas (las filas de la matriz de Jacobi son componentes de los gradientes $\nabla f_1, \ldots, \nabla f_m$).

De acuerdo con los resultados del párrafo anterior, esto significa que las funciones del sistema (43.1) son independientes en cierto entorno del punto $x^{(0)}$. Por cuanto en un espacio n-dimensional no pueden haber más de n vectores linealmente independientes y el rango de la matriz no puede exceder del número de columnas, entonces, de la condición 2) proviene que $m \leqslant n$.

Conforme a la condición 2), en el punto $x^{(0)}$ al menos uno de los determinantes del tipo

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(x_{i_1}, \ldots, x_{i_m})}$$

es distinto de cero. Supongamos, para concretar, que en el punto $x^{(0)}$

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_m)} \neq 0. \tag{43.6}$$

Entonces, en virtud del teorema sobre las funciones implícitas (véase el p. 41.3), el sistema de ecuaciones (43.3) en cierto entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ puede resolverse respecto de las variables $x_1, \ldots, x_m$:

$$\begin{aligned} x_1 &= \varphi_1(x_{m+1}, \ldots, x_n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_m &= \varphi_m(x_{m+1}, \ldots, x_n) \end{aligned} \tag{43.7}$$

Al sustituir los valores de $x_1, \ldots, x_m$, proporcionados por las fórmulas (43.7), en $y = f_0(x)$, es decir, al examinar la composición de las funciones f_0 y $\varphi_1, \ldots, \varphi_m$, obtendremos una función

$$y = f_0(\varphi_1(x_{m+1}, \ldots x_n), \ldots, \varphi_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) \stackrel{\text{def}}{=} g(x_{m+1}, \ldots, x_n) \tag{43.8}$$

de $n-m$ variables $x_{m+1}, \ldots, x_n$ que está definida y es continuamente derivable en cierto entorno del punto $\tilde{x}^{(0)} = (x_{m+1}^{(0)}, \ldots, x_n^{(0)})$ en el espacio $(n-m)$-dimensional R^{n-m}.

Como, de acuerdo con el teorema sobre las funciones implícitas, las condiciones (43.3) y (43.7) son equivalentes, resulta válida la siguiente afirmación.

El punto $x^{(0)}$ es un punto de extremo condicionado (estricto) para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3) *cuando, y sólo cuando, $\tilde{x}^{(0)}$ sea un punto de extremo ordinario (estricto) de la función* (43.8).

Si $\tilde{x}^{(0)}$ es un punto de extremo ordinario de la función g, será el punto estacionario de esta función (véase el p. 40.1):

$$dg(\tilde{x}^{(0)}) = 0. \tag{43.9}$$

Recordemos que la diferencial es una función homogénea lineal y el hecho de que ella es nula significa que es nula dicha función, cualesquiera que sean los valores de sus argumentos, en el caso dado, para cualesquiera $dx_{m+1}, dx_{m+2}, \ldots, dx_n$. Esto es factible, evidentemente, cuando y sólo cuando, todos los coeficientes de los

argumentos citados, es decir, las derivadas $\dfrac{\partial g}{\partial x_{m+k}}$, $k = 1, 2, \ldots n - m$, se anulan en el punto $x^{(0)}$. La condición (43.9) es necesaria para que en el punto $x^{(0)}$ haya un extremo condicionado.

De este modo, el método, basado en la resolución del sistema de ecuaciones (43.3), permite reducir el problema de búsqueda del extremo condicionado al problema, ya estudiado, sobre un extremo ordinario. Precisamente este procedimiento se ha empleado en los ejemplos considerados más arriba. Sin embargo, resulta a menudo imposible o muy embarazoso expresar la solución del sistema (43.3) en términos de las funciones elementales, razón por la cual es deseable disponer de un método que permita hallar el extremo condicionado sin resolver el sistema (43.3). Tal método va expuesto más abajo.

43.2. MÉTODO DE LOS MULTIPLICADORES DE LAGRANGE PARA BUSCAR LOS PUNTOS DE EXTREMO CONDICIONADO

En este punto se supondrá que todas las funciones $f_0, f_1, \ldots, f_m$ son continuamente derivables en el conjunto abierto G.

Teorema 1. *Sea $x^{(0)}$ un punto de extremo condicionado de la función f_0, cumpliéndose las ecuaciones de enlace* (43.3). *En este punto los gradientes ∇f_0, $\nabla f_1, \ldots, \nabla f_m$ son linealmente dependientes, es decir, existen tales números λ_0, $\lambda_1, \ldots, \lambda_m$, de los cuales no todos son nulos, que*

$$\lambda_0 \nabla f_0 + \lambda_1 \nabla f_1 + \ldots + \lambda_m \nabla f_m = 0. \tag{43.10}$$

Corolario. *Si en el punto $x^{(0)}$ de extremo condicionado de la función f_0 respecto de las ecuaciones de enlace* (43.3) *los gradientes $\nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, es decir, el rango de la matriz de Jacobi*

$$\left(\frac{\partial f_j}{\partial x_i}\right), j = 1, 2, \ldots, m \quad i = 1, 2, \ldots, n,$$

es igual a m, entonces existen tales $\lambda_1, \ldots, \lambda_m$ que en dicho punto

$$\nabla f_0 + \sum_{j=1}^{m} \lambda_j \nabla f_j = 0, \tag{43.11}$$

es decir, ∇f_0 es una combinación lineal de los gradientes $\nabla f_1, \ldots, \nabla f_m$.

En la forma coordenada esta condición tiene por expresión: para cualquier $i = 1, 2, \ldots, n$ en el punto $x^{(0)}$

$$\frac{\partial f_0}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial f_j}{\partial x_i} = 0. \tag{43.12}$$

La función

$$F(x) \stackrel{\text{def}}{=} f_0(x) + \sum_{j=1}^{m} \lambda_j f_j(x), \tag{43.13}$$

donde los números $\lambda_1, \ldots, \lambda_m$ satisfacen la condición (43.12), se llama *función de Lagrange* del problema en consideración y los propios números, *multiplicadores de Lagrange*.

La condición (43.12) significa que si $x^{(0)}$ es un punto de extremo condicionado de la función f_0 respecto de la ecuación de enlace (43.3), será punto estacionario para la función de Lagrange, es decir,

$$\frac{\partial F(x^{(0)})}{\partial x_i} = 0, i = 1, 2, \ldots, n. \tag{43.14}$$

Antes de demostrar el teorema aclaremos su sentido y mostremos cómo usarlo para la obtención de los puntos de extremo condicionado. Ante todo fijemos la atención en que en la función del tipo (43.13), para los números arbitrarios $\lambda_1, \ldots, \lambda_m$, cualquier punto de su extremo condicionado es a la vez el punto de extremo condicionado de la función de partida f_0, y viceversa. Elijamos tales valores de $\lambda_1, \ldots, \lambda_n$ que se cumplan las condiciones (43.12), es decir, que el punto dado de extremo condicionado resulte ser también el punto estacionario de la función (43.11).

Para la obtención de los puntos de extremo condicionado, conviene considerar un sistema de $n + m$ ecuaciones (43.3) y (43.10) respecto de las incógnitas $x_1^{(0)}, \ldots, x_n^{(0)}, \lambda_1, \ldots, \lambda_m$ y resolverlo (si esto es factible), hallando $x_1^{(0)}, \ldots, x_n^{(0)}$ y eliminando, cuanto sea posible, $\lambda_1, \ldots, \lambda_m$. El teorema enunciado afirma que todos los puntos de extremo condicionado estarán entre los puntos $(x_1^{(0)}, \ldots, x_n^{(0)})$, determinados de este modo. La cuestión de qué puntos entre los mencionados serán realmente puntos de extremo condicionado requiere unas investigaciones adicionales; algunos razonamientos referentes a esta cuestión serán aducidos en el p. 43.5*.

DEMOSTRACIÓN DEL TEOREMA. Demostremos una afirmación que es equivalente al teorema: si en el punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ que satisface las ecuaciones de enlace

$$f_k(x^{(0)}) = 0, k = 1, 2, \ldots, m, \tag{43.15}$$

los gradientes $\nabla f_0, \nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, entonces $x^{(0)}$ no es el punto de extremo condicionado.

Así pues, sean $\nabla f_0, \nabla f_1, \ldots, \nabla f_m$ linealmente independientes y, por consiguiente, el rango de la matriz de Jacobi $\left(\frac{\partial f_i}{\partial x_j}\right), j = 0, 1, \ldots, m, i = 1, 2, \ldots, n$, es igual a $m + 1$. Entonces, en dicha matriz existe un menor de orden $m + 1$ distinto de cero. Para concretar, convengamos en considerar que el menor se ha formado por las primeras $m + 1$ columnas, es decir,

$$\left.\frac{\partial(f_0, f_1, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_{m+1})}\right|_{x = x^{(0)}} \neq 0. \tag{43.16}$$

El conjunto G es abierto, por lo cual existe tal $\delta_0 > 0$, que para todo δ, $0 < \delta < \delta_0$, el cubo

$$Q^n_\delta = \{x : |x_i - x_i^{(0)}| < \delta, i = 1, 2, \ldots, n\}$$

se dispone dentro de G y, por ende, en este cubo están definidas todas las funciones $f_0, f_1, \ldots, f_m$.

Fijemos $x_{m+2} = x_{m+2}^{(0)}, \ldots, x_n = x_n^{(0)}$ e introduzcamos las designaciones

$$x^* = (x_1, \ldots, x_{m+1})$$

$$Q_\delta^{m+1} = \{x^* : |x_i - x_i^{(0)}| < \delta, i = 1, 2, \ldots, m+1\}.$$

Obviamente, las funciones $f_j(x_1, \ldots, x_{m+1}, x_{m+2}^{(0)}, \ldots, x_n^{(0)})$, $j = 0, 1, \ldots, m$, están definidas y son continuamente derivables en todo punto de Q_δ^{m+1}. Consideraremos una aplicación $\Phi: Q_\delta^{m+1} \to R^{m+1}$, definida por las fórmulas

$$\begin{aligned} y_1 &= f_0(x_1, \ldots, x_{m+1}, x_{m+2}^{(0)}, \ldots, x_n^{(0)}) \\ y_2 &= f_1(x_1, \ldots, x_{m+1}, x_{m+2}^{(0)}, \ldots, x_n^{(0)}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{m+1} &= f_{m+1}(x_1, \ldots, x_{m+1}, x_{m+2}^{(0)}, \ldots, x_n^{(0)}). \end{aligned} \tag{43.17}$$

En vista de (43.16), para el punto $x^{*(0)} = (x_1^{(0)}, \ldots, x_{m+1}^{(0)})$ tenemos

$$\left.\frac{\partial(y_1, \ldots, y_{m+1})}{\partial(x_1, \ldots, x_{m+1})}\right|_{x^* = x^{*(0)}} = \left.\frac{\partial(f_0, f_1, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_{m+1})}\right|_{x = x^{(0)}} \neq 0,$$

y debido a (43.15),

$$\Phi(x^{*(0)}) = (f_0(x^{(0)}), 0, \ldots, 0).$$

Por ello, (véase el teorema 7 en el p. 41.8 sobre la invertibilidad local de la aplicación continuamente derivable en un punto, donde su jacobiano es distinto de cero) existe tal $\varepsilon > 0$ que en el entorno

$$V = \{y = (y_1, \ldots, y_{m+1}) : |y_1 - f_0(x^{(0)})| < \varepsilon, |y_j| < \varepsilon, j = 2, 3, \ldots, m+1\}$$

(fig. 171) está definida una aplicación inversa respecto de Φ, y, por lo tanto, en cualquier punto de este entorno se aplica algún punto de Q_δ^{m+1}.

En particular, puesto que para cualquier η, $0 < \eta < \varepsilon$, tiene lugar la inclusión $(f(x^{(0)}) \pm \eta, 0, \ldots, 0) \in V$, entonces en el cubo Q_δ^{m+1} existen los puntos $x'^* = (x_1', \ldots, x_{m+1}')$ y $x''^* = (x_1'', \ldots, x_{m+1}'')$ que, realizándose la aplicación Φ, se aplican en los puntos indicados del entorno V:

$$\Phi(x'^*) = (f(x^{(0)}) + \eta, 0, \ldots, 0), \Phi(x''^*) = (f(x^{(0)}) - \eta, 0, \ldots, 0).$$

Si ponemos, para abreviar, $x' = (x_1', \ldots, x_{m+1}', x_{m+2}^{(0)}, \ldots, x_n^{(0)})$ y $x'' = (x_1'', \ldots, x_{m+1}'', x_{m+2}^{(0)}, \ldots, x_n^{(0)})$, entonces en la inscripción coordenada obtendremos (véase (43.17)):

$$f_0(x') = f_0(x^{(0)}) + \eta > f(x^{(0)}),$$

$$f_k(x') = 0, k = 1, 2, \ldots, m, x' \in Q_\delta^\eta$$

y

$$f_0(x'') = f_0(x^{(0)}) - \eta < f(x^{(0)}),$$

$$f_k(x'') = 0, k = 1, 2, \ldots, m, x'' \in Q_\delta^\eta.$$

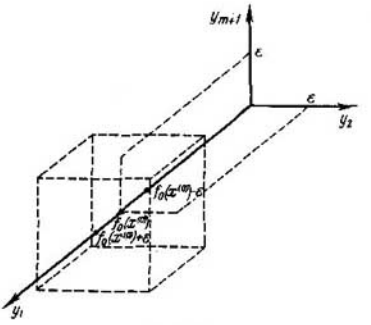


Fig. 171

Puesto que $\delta > 0$, $0 < \delta < \delta_0$, es arbitrario, esto es indicio de que $x^{(0)}$ no es el punto de extremo condicionado. □

DEMOSTRACIÓN DEL COROLARIO. Si los vectores $\nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, en la igualdad (43.10) tenemos $\lambda_0 \neq 0$, puesto que si fuera $\lambda_0 = 0$, los vectores mencionados serían, en virtud de (43.10), linealmente dependientes. Al dividir ambos miembros de (43.10) por λ_0, obtendremos una igualdad del tipo (43.11). □

43.3°. INTERPRETACIÓN GEOMÉTRICA DEL MÉTODO DE LAGRANGE

Daremos ahora algunas explicaciones geométricas al teorema 1. Consideraremos, para no complicar, el caso de un extremo condicionado de la función de dos variables $z = f(x, y)$ verificándose la ecuación de enlace $\varphi(x, y) = 0$.

Supongamos que las funciones f y φ son continuamente derivables en el entorno del punto (x_0, y_0), $\nabla\varphi(x_0, y_0) = \left(\frac{\partial\varphi(x_0, y_0)}{\partial x}, \frac{\partial\varphi(x_0, y_0)}{\partial y}\right) \neq 0$ y $\varphi(x_0, y_0) = 0$. En vista de la condición $\nabla\varphi(x_0, y_0) \neq 0$, de acuerdo con el teorema sobre las funciones implícitas, la ecuación $\varphi(x, y) = 0$ en el entorno del punto (x_0, y_0) define cierta curva suave cuya representación explícita tiene por expresión o bien $y = y(x)$, o bien $x = x(y)$. Por cuanto nos interesan sólo los puntos suficientemente próximos a (x_0, y_0), la curva mencionada se llamará simplemente curva $\varphi(x, y) = 0$ (es decir, en otras palabras, en lo que sigue se considerará siempre la restricción de las funciones f y φ al entorno indicado del punto (x_0, y_0)).

El gradiente $\nabla\varphi(x_0, y_0)$ es una normal a la curva $\varphi(x, y) = 0$ en el punto (x_0, y_0) (p. 20.6). Designemos con τ el vector unidad tangente a la curva $\varphi(x, y)$ en el punto (x_0, y_0). Supongamos, para concretar, que la curva en consideración se da por la

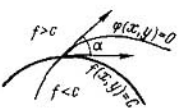


Fig. 172

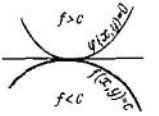


Fig. 173

ecuación $y = y(x)$. Si el punto (x_0, y_0) es un punto de extremo condicionado, entonces x_0 es un punto de extremo ordinario para la funcion $g(x) \stackrel{\text{def}}{=} f(x, g(x))$ (véase el p. 43.1), y, por lo tanto, $g'(x) = 0$, es decir, la derivada de la función f en el punto (x_0, y_0) según la dirección de la curva $\varphi(x, y) = 0$, o bien, que es lo mismo (vease el p. 20.7), según la dirección del vector τ, es igual a cero

$$\frac{\partial f(x_0, y_0)}{\partial \tau} = (\nabla f(x_0, y_0), \tau) = 0.$$

Esto significa la ortogonalidad del gradiente $\nabla f(x_0, y_0)$ y del vector tangente τ, lo que es equivalente al carácter colineal de los vectores $\nabla f(x_0, y_0)$ y $\nabla \varphi(x_0, y_0)$:

$$\nabla f(x_0, y_0) = \lambda \nabla \varphi(x_0, y_0),$$

es decir, se cumple la condición (43.11). El cumplimiento de esta condición en el punto de extremo condicionado puede explicarse también de otro modo. Sea $f(x_0, y_0) \stackrel{\text{def}}{=} c$. Si en el punto (x_0, y_0) no se cumple la condición (43.11), es decir, cuando los gradientes ∇f y $\nabla \varphi$ no son colineales, esto significa que en dicho punto $\nabla f \neq 0$ y la línea de nivel $f(x, y) = c$ corta en este punto la curva $\varphi(x, y) = 0$, haciendo con ésta cierto ángulo α, distinto de 0 y π (fig. 172). Por eso en todo entorno suficientemente pequeño del punto (x_0, y_0) una parte de la curva $\varphi(x, y) = 0$ resultará ser dispuesta en el dominio $f < c$ ("en el dominio de los valores menores") y otra parte, en el dominio $f > c$ ("en el dominio de los valores mayores"). Esto significa que en el punto (x_0, y_0) no hay extremo condicionado de que se trata.

En el caso en que los vectores ∇f y $\nabla \varphi$ sean colineales, $\nabla f = \lambda \nabla \varphi$, una parte de la curva $\varphi(x, y) = 0$ puede pertenecer a cierto entorno del punto (x_0, y_0), íntegramente dispuesto en el dominio de valores menores $f < c$ (fig. 173) o bien en el dominio de valores mayores $f > c$. En este caso en el punto (x_0, y_0) se alcanza un extremo condicionado.

Sin embargo, siendo colineales los vectores ∇f y $\nabla \varphi$ la curva $\varphi(x, y) = 0$ puede también disponerse en cualquier entorno suficientemente pequeño del punto (x_0, y_0), con la particularidad de que una parte de dicha curva se dispondrá en el dominio de valores menores y la otra parte, en el de valores mayores de la función f (fig. 174); en este caso en el punto (x_0, y_0) tampoco habrá extremo condicionado. Una situación semejante aparece, por ejemplo, cuando las curvas $f(x, y) = c$ y $\varphi(x, y) = 0$ tienen en el punto (x_0, y_0) una tangente común, con la particularidad de que la curva $f(x, y) = c$ está dispuesta en un entorno suficientemente pequeño del punto (x_0, y_0) por un lado de la tangente mencionada, mientras que la curva $\varphi(x, y) = 0$ sufre inflexión en dicho punto, pasando de un lado de la tangente al otro lado.

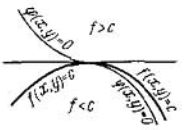


Fig. 174

Lo dicho aclara aquella circunstancia que (43.10) es una *condición necesaria* para que haya extremo condicionado, *pero no suficiente*.

Los razonamientos geométricos aducidos referentes a la cuestión acerca del extremo condicionado quedan válidos también en el caso multidimensional.

43.4*. PUNTOS ESTACIONARIOS DE LA FUNCIÓN DE LAGRANGE

Se dará aquí la descripción de los puntos estacionarios de la función de Lagrange (43.13) por medio de la función $g(x_{m+1}, \ldots, x_n)$ introducida en el p. 43.1 (véase (43.8)). Demostremos previamente un lema sencillo conocido del curso de álgebra lineal.

Sea dado un sistema de ecuaciones homogéneas lineales

$$a_{i1}x_1 + \ldots + a_{in}x_n = 0, i = 1, 2, \ldots, m, \tag{43.18}$$

y una ecuación homogénea lineal más

$$b_1x_1 + \ldots + b_nx_n = 0. \tag{43.19}$$

El sistema de ecuaciones obtenido por adición de la ecuación (43.19) al sistema (43.18) se denominará *sistema ampliado* (43.18)—(43.19).

Lema. *Para que el sistema ampliado* (43.18) — (43.19) *sea equivalente al sistema básico* (43.18), *es necesario y suficiente que la ecuación* (43.19) *sea una combinación lineal del sistema de ecuaciones* (43.18).

Corolario. *Para que la ecuación* (43.19) *sea una combinación lineal de las ecuaciones* (43.18) *o bien, que es lo mismo, para que el vector*

$$b \stackrel{\text{def}}{=} (b_1, \ldots, b_n) \tag{43.20}$$

constituya una combinación lineal de los vectores

$$a_i \stackrel{\text{def}}{=} (a_{i1}, \ldots, a_{in}), i = 1, 2, \ldots, m, \tag{43.21}$$

es necesario y suficiente que toda solución del sistema (43.18) *sea la solución de la ecuación* (43.19).

DEMOSTRACIÓN DEL LEMA. Supongamos que el rango de la matriz (a_{ij}) de los coeficientes del sistema (43.18) es igual a m_0. Es evidente que $m_0 \leqslant m$. Si $m_0 < m$, entonces $m - m_0$ ecuaciones del sistema (43.18) son combinaciones lineales de las demás. Al desechar aquellas $m - m_0$ ecuaciones lineales que son combinaciones lineales de las restantes, obtendremos un sistema de m_0 ecuaciones linealmente independientes equivalente al sistema (43.18), con la particularidad de que la ecuación

(43.19) constituye una combinación lineal de las ecuaciones del sistema (43.18) cuando, y sólo cuando, es una combinación lineal del sistema mencionado formado por m_0 ecuaciones restantes. Por eso, desde el mismo principio convengamos en considerar que $m = m_0$, es decir, que el rango de la matriz (a_{ij}) de los coeficientes del sistema (43.18) es igual a m, es decir, al número de ecuaciones de este sistema.

Supongamos que los sistemas (43.18) y (43.18) — (43.19) son equivalentes. Esto significa que los espacios de sus soluciones coinciden. Por cuanto todas las ecuaciones del sistema básico (43.18) figuran en el sistema ampliado (43.18) — (43.19), toda solución del sistema ampliado será también la solución del sistema básico, es decir, el espacio de soluciones del sistema ampliado está contenido en el espacio de soluciones del sistema básico. Por consiguiente, la coincidencia de estos espacios equivale a la igualdad de sus dimensiones.

La dimensión s del espacio de soluciones de un sistema de ecuaciones homogéneas lineales es igual, como se sabe, al número de las incógnitas n de dicho sistema, del cual se ha sustraído el rango r de la matriz de los coeficientes del sistema: $s = n - r$. De aquí se deduce que la equivalencia de los sistemas (43.18) y (43.18)—(43.19) implica la igualdad de los rangos de sus matrices. El rango de la matriz de los coeficientes del sistema (43.18) es igual, por hipótesis, a m, es decir, los vectores (43.21) son linealmente independientes.

El rango de la matriz de los coeficientes del sistema ampliado (48.18) — (43.19), de acuerdo con lo expuesto en nuestras condiciones, es también igual a m. Por ello, los vectores (véanse (43.20) y (43.21))

$$b, a_1, \ldots, a_m \tag{43.22}$$

son linealmente dependientes. Esto es indicio de que b es una combinación lineal de los vectores $a_1, \ldots, a_m$.

Efectivamente, la dependencia lineal de los vectores (43.22) significa que existen tales números $\mu_0, \mu_1, \ldots, \mu_m$, no todos nulos, que

$$\mu_0 b + \mu_1 a_1 + \ldots + \mu_m a_m = 0. \tag{43.23}$$

Aquí, a ciencia cierta, $\mu_0 \neq 0$, puesto que de lo contrario los vectores $a_1, \ldots, a_m$ resultarían ser linealmente dependientes. Al dividir la igualdad (43.23) por μ_0, llegamos a que b es una combinación lineal de los vectores $a_1, \ldots, a_m$.

Viceversa, si b es una combinación lineal de los vectores (43.21), entonces los sistemas de vectores (43.21) y (43.22) tienen cada uno exactamente m vectores linealmente independientes, es decir, los rangos de las matrices de los coeficientes de los sistemas de ecuaciones (43.18) y (43.18) — (43.19) son iguales.

Así pues, la condición de que el vector b es una combinación lineal de los vectores (43.21):

$$b = \lambda_1 a_1 + \ldots + \lambda_m a_m$$

es equivalente a la igualdad de rangos de las matrices de los coeficientes de los sistemas básico y ampliado en consideración y, por consiguiente, significa su equivalencia. □

La afirmación del corolario proviene seguidamente del lema, ya que los sistemas (43.18) y (43.18) — (43.19) son, evidentemente equivalentes cuando, y sólo cuando,

toda solución del sistema (43.18) es, a la vez, la solución del sistema (43.19); las demás ecuaciones de estos sistemas son simplemente coincidentes. □

OBSERVACIÓN 1. El lema demostrado y sus corolarios tienen una sencilla interpretación geométrica en el espacio vectorial *euclídeo* n-dimensional R^n, es decir, en un espacio n-dimensional provisto de un producto escalar. Haciendo uso de la designación del producto escalar, el sistema (43.18) puede escribirse en la forma

$$(a_i, x) = 0, i = 1, 2, \ldots, m, \tag{43.24}$$

y la ecuación (43.19), en la forma

$$(b, x) = 0, \tag{43.25}$$

donde los vectores $a_1, \ldots, a_m$ y b están definidos en (43.20) y (43.21), mientras que $x = (x_1, \ldots, x_n)$.

El conjunto de toda una serie de combinaciones lineales de los vectores $a_1, \ldots, a_m$ forma un subespacio del espacio R^n y se denomina *subespacio tendido sobre dichos vectores*. Designémoslo $\mathcal{L}(a_1, \ldots, a_m)$.

El conjunto de soluciones del sistema (43.24) se compone de todos los vectores x ortogonales al subespacio $\mathcal{L}(a_1, \ldots, a_m)$. Designemos este conjunto de soluciones con T. Es también un subespacio del espacio R^n.

Los subespacios $L \stackrel{\text{def}}{=} \mathcal{L}(a_1, \ldots, a_m)$ y T se denominan *complementos ortogonales* uno al otro en el *espacio* R^n.

Dado que $L = \mathcal{L}(a_1, \ldots, a_m)$, la representación del vector b en forma de una combinación lineal de los vectores $a_1, \ldots, a_m$ es equivalente a su pertenencia al subespacio L del espacio R^n: $b \in L$. Esta condición es equivalente, a su vez, a la ortogonalidad del vector b al subespacio T: $b \perp T$, la que significa que para todo $x \in T$ tiene lugar la igualdad $(b, x) = 0$, es decir, que toda solución x del sistema (43.24) es la solución de la ecuación (43.25). Esto es precisamente la afirmación del corolario del lema.

OBSERVACIÓN 2. Recordemos un método por medio del cual pueden obtenerse todas las soluciones de un sistema homogéneo de ecuaciones lineales. Supongamos que el sistema (43.18) se compone de las ecuaciones linealmente independientes. En este caso el rango de la matriz de sus coeficientes es igual a m. Esto significa que existe un menor de dicha matriz de orden m, distinto de cero. Sea, para concretar,

$$\begin{vmatrix} a_{11} \ldots a_{1m} \\ a_{m1} \ldots a_{mm} \end{vmatrix} \neq 0. \tag{43.26}$$

En este caso todas las soluciones del sistema (43.18) pueden obtenerse, fijando arbitrariamente las últimas $n - m$ coordenadas del vector $(x_1, \ldots, x_n)$. Las coordenadas restantes se obtienen unívocamente del sistema de ecuaciones (43.18). En efecto, tomemos una solución arbitraria $(x_1^{(0)}, \ldots, x_n^{(0)})$ del sistema (43.18). Realizada la sustitución $x_{m+1} = x_{m+1}^{(0)}, \ldots, x_n = x_n^{(0)}$ en (43.18), se obtendrá un sistema de m ecuaciones lineales (con m incógnitas $x_1, \ldots, x_m$) tal que la matriz de los coeficientes de este sistema es regular, en virtud de la condición (43.26). Por esta razón existen los únicos valores $x_1, \ldots, x_m$ que satisfacen el sistema obtenido. Por cuanto $(x_1^{(0)}, \ldots, x_n^{(0)})$ fue también la solución del sistema (43.18), se tiene $x_1 = x_1^{(0)}, \ldots, x_m = x_m^{(0)}$.

Pasemos ahora al análisis de los puntos estacionarios de la función de Lagrange.

Teorema 2. *Supongamos que las funciones* $f_0, f_1, \ldots, f_m$ *son continuamente derivables en la región* $G \subset R^n$, $x^{(0)} \in G$,

$$f_i(x^{(0)}) = 0,\ i = 1, 2, \ldots, m,$$

y el rango de la matriz de Jacobi de las funciones $f_1, \ldots, f_m$ *en el punto* $x^{(0)}$ *es igual a* m. *Para que en el punto* $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ *el gradiente* ∇f_0 *sea una combinación lineal de los gradientes* $\nabla f_1, \ldots, \nabla f_m$ *es necesario y suficiente que el punto* $x^{(0)} = (x_{m+1}^{(0)}, \ldots, x_n^{(0)})$ *sea estacionario para la funcion* $g(x) = g(x_{m+1}, \ldots, x_n)$ *(véase (43.8)).*

Recordemos que si en el punto $x^{(0)}$ el gradiente ∇f_0 es una combinación lineal

$$\nabla f_0 = \lambda_1 \nabla f_1 + \ldots + \lambda_m \nabla f_m, \tag{43.27}$$

de los gradientes $\nabla f_1, \ldots, \nabla f_m$, esto es equivalente a que existe una función de Lagrange

$$F = f_0 - \lambda_1 f_1 - \ldots - \lambda_m f_m, \tag{43.28}$$

para la cual el punto $x^{(0)}$ es estacionario:

$$\frac{\partial F(x^{(0)})}{\partial x_i} = 0,\ i = 1, 2, \ldots, n. \tag{43.29}$$

Esta es simplemente una inscripción coordenada de la condición (43.27), pues, en virtud de (43.28),

$$\frac{\partial F}{\partial x_i} = \frac{\partial f_0}{\partial x_i} - \lambda_1 \frac{\partial f_1}{\partial x_i} - \ldots - \lambda_m \frac{\partial f_m}{\partial x_i},\ i = 1, \ldots, m.$$

DEMOSTRACIÓN. Por hipótesis, el rango de la matriz de Jacobi del sistema de funciones $f_1, \ldots, f_m$ en el punto $x^{(0)}$ es igual a m. Convengamos en considerar, para concretar, como en el p. 43.1, que

$$\left.\frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_m)}\right|_{x^{(0)}} \neq 0. \tag{43.30}$$

Sustituyamos en la ecuación de enlace (43.3) las funciones (43.7) que representan las soluciones de estas ecuaciones y derivemos las identidades obtenidas respecto de las variables $x_{m+1}, \ldots, x_n$. Obtendremos para el punto $x^{(0)}$ las igualdades $df_i(x^{(0)}) = 0$, $i = 1, 2, \ldots, m$, válidas para cualesquiera incrementos $dx_{m+1}, \ldots, dx_n$ de las variables independientes $x_{m+1}, \ldots, x_n$ (recordemos que la diferencial es una función lineal definida en todo el espacio). Al hacer uso de la invariación de la forma de la primera diferencial respecto de la elección de las variables, llegamos a que en el punto $x^{(0)}$ se cumplen las igualdades

$$\frac{\partial f_i}{\partial x_1} dx_1 + \ldots + \frac{\partial f_i}{\partial x_m} dx_m + \frac{\partial f_i}{\partial x_{m+1}} dx_{m+1} + \ldots$$

$$\ldots + \frac{\partial f_i}{\partial x_n} dx_n = 0,\ i = 1, 2, \ldots, m, \tag{43.31}$$

donde $dx_{m+1}, \ldots dx_n$ son arbitrarios, mientras que $dx_1, \ldots, dx_m$ se hallan a partir de las fórmulas (43.7). De este modo, el vector

$$dx = (dx_1, \ldots, dx_m, dx_{m+1}, \ldots, dx_n) \tag{43.32}$$

es la solución del sistema homogéneo lineal (43.31).

Cabe notar que en virtud de la condición (43.30), los valores de $dx_1, \ldots, dx_m$, para $dx_{m+1}, \ldots, dx_n$ dados, se hallan también unívocamente del sistema (43.31). De la observación 2 se infiere, además, que mediante el método indicado se obtienen todas las soluciones del sistema (43.31).

El carácter estacionario del punto $x^{(0)}$ para la función $g(x) = g(x_{m+1}, \ldots, x_n)$ significa que $dg(x^{(0)}) = 0$. Por ser invariante la forma de la primera diferencial, esta igualdad puede escribirse más detalladamente como sigue

$$\frac{\partial f_0}{\partial x_1} dx_1 + \ldots + \frac{\partial f_0}{\partial x_m} dx_m + \frac{\partial f_0}{\partial x_{m+1}} dx_{m+1} + \ldots + \frac{\partial f_0}{\partial x_n} dx_n = 0, \tag{43.33}$$

donde $dx_{m+1}, \ldots, dx_n$ pueden prefijarse de modo arbitrario, mientras que $dx_1, \ldots, dx_m$ se deben hallar de las fórmulas (43.7), o bien, que proporciona el mismo resultado, de las fórmulas (43.31). En otras palabras, cualquier solución del sistema de ecuaciones (43.31) es, a la vez, la solución de la ecuación (43.33). Conforme al corolario del lema, esto resulta posible cuando, y sólo cuando, la ecuación (43.33) es una combinación lineal de las ecuaciones del sistema (43.31), es decir, cuando existen tales números $\lambda_1, \ldots, \lambda_m$ que

$$\nabla f_0 = \lambda_1 \nabla f_1 + \ldots + \lambda_m \nabla f_m. \quad \square$$

OBSERVACIÓN 3. De acuerdo con la observación 2, la totalidad de todas las soluciones del sistema de ecuaciones (43.31) forma un subespacio T del espacio R^n que es un *complemento ortogonal* al subespacio $L = \mathcal{L}(\nabla f_1, \ldots, \nabla f_m)$. Todo vector $y \in T$ es ortogonal respecto de cualquier gradiente ∇f_i, por lo cual será natural llamarlo *vector tangente* en el punto $x^{(0)}$ a la hipersuperficie $f_i(x) = 0$, la cual representa un conjunto de nivel (véase § 19) de la función f_i, $i = 1, \ldots, m$.

De este modo, el espacio de soluciones T del sistema (43.31) se compone de los vectores tangentes simultáneamente a todas las hipersuperficies $f_i(x) = 0$, $i = 1, \ldots, m$, y por esta razón lo llaman espacio tangente de la intersección de todas las hipersuperficies $f_i(x) = 0$, $i = 1, 2, \ldots, m$. Recordemos que los vectores del espacio tangente T, es decir, las soluciones del sistema (43.31) han sido designados mediante dx (43.32).

Puesto que, de acuerdo con el teorema 2, en el punto de extremo condicionado tiene lugar la inclusión

$$\nabla f_0 \in L = \mathcal{L}(\nabla f_1, \ldots, \nabla f_m),$$

se tiene

$$\nabla f_0 \perp T.$$

En otras palabras, el gradiente ∇f_0 es ortogonal simultáneamente a todas las tangentes dx a las superficies $f_i(x) = 0$, $i = 1, 2, \ldots, m$:

$$(\nabla f_0, dx) = 0$$

(ésta es otra notación de la ecuación (43.33)), es decir, el gradiente ∇f_0 es perpendicular al espacio tangente T en el punto $x^{(0)}$. Pero el conjunto de todos los vectores ortogonales a ∇f_0 forma un subespacio $(n-1)$-dimensional T_0, denominado espacio tangente a la hipersuperficie $f_0(x) = f_0(x^{(0)})$. En virtud de lo dicho anteriormente, cada vector de T, siendo ortogonal al gradiente ∇f_0, pertenece a T_0, es decir, $T \subset T_0$.

Así pues, *si $x^{(0)}$ es un punto de extremo condicionado, entonces $T \subset T_0$, es decir, el espacio tangente en el punto $x^{(0)}$ de intersección de todas las hipersuperficies, definidas por las ecuaciones de enlace, está contenido en el espacio tangente en el mismo punto de la hipersuperficie $f_0(x) = f_0(x^{(0)})$.*

OBSERVACIÓN 4. Del teorema 2 se deduce una vez más el corolario del teorema 1. Efectivamente, si $x^{(0)}$ es un punto de extremo condicionado, entonces $x^{(0)}$ será el punto de extremo ordinario para la función g (véase el p. 43.1) y, por consiguiente, su punto estacionario. Por ello, de acuerdo con el teorema 2, $x^{(0)}$ es el punto estacionario para la función de Lagrange, es decir, se cumple la condición (43.29).

43.5*. CONDICIONES SUFICIENTES PARA LOS PUNTOS DE EXTREMO CONDICIONADO

En este punto se supondrán también vigentes todas las suposiciones impuestas en las funciones f_0 y f_i, $i = 1, 2, \ldots, m$, en el p. 43.1. Sea

$$F = f_0 + \sum_{i=1}^{m} \lambda_i f_i$$

una función de Lagrange (véase (43.13)) para la función f_0 y las ecuaciones de enlace (43.3). Supongamos que $x^{(0)} \in G$ satisface las ecuaciones de enlace (43.3) y es un punto estacionario de la función de Lagrange, es decir, un punto cuyas coordenadas satisfacen el sistema de ecuaciones (43.12) y (43.3). Nuestro objeto es la obtención de un método, por cuyo intermedio se pueden establecer las condiciones suficientes para que $x^{(0)}$ sea punto de extremo condicionado del problema en consideración.

Observemos ante todo que si el punto $x \in G$ satisface las ecuaciones de enlace (43.3), entonces

$$\Delta f_0 = f_0(x) - f_0(x^{(0)}) = F(x) - F(x^{(0)}) = \Delta F. \tag{43.34}$$

De aquí se ve en seguida que si $x^{(0)}$ es un punto de extremo habitual para la función F, es decir, si ΔF no cambia de signo en cierto entorno del punto $x^{(0)}$, entonces $x^{(0)}$ es el punto de extremo condicionado para la función f_0.

En efecto, de (43.34) se deduce en este caso que el incremento Δf_0 para los valores admisibles de x, es decir, para aquellos que satisfacen las ecuaciones de enlace, tampoco cambia de signo. Esta condición suficiente impone, sin embargo, una restricción demasiado fuerte sobre el comportamiento de la función de Lagrange $F(x)$ en el punto considerado: ésta debe tener extremo ordinario, lo que hace muy estrecho el dominio de la aplicación eventual de la condición indicada en la resolución de los problemas. Por esta razón parece conveniente obtener un criterio suficiente más general para un extremo condicionado.

Supongamos que $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ satisface las ecuaciones de enlace (43.3). Volvamos a considerar la función (43.8), es decir, una función $g(\xi) = g(x_{m+1}, \ldots, x_n)$, $\xi = (x_{m+1}, \ldots, x_n)$, que se obtiene de $f_0(x) = f_0(x_1, \ldots, x_n)$ a condición de que $x_1, \ldots, x_m$ son funciones de las variables $x_{m+1}, \ldots, x_n$, determinadas por las ecuaciones de enlace (43.3) en cierto entorno del punto $x^{(0)}$. Supondremos adicionalmente que $f_0(x)$ y $f_i(x)$, $i = 1, 2, \ldots, m$, son dos veces continuamente derivables en el punto $x^{(0)}$.

Se ha notado más arriba (véase el p. 43.1) que $x^{(0)}$ es un punto de extremo condicionado (estricto) para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3) cuando, y sólo cuando, $\xi^{(0)} = (x_{m+1}^{(0)}, \ldots, x_n^{(0)})$ es un punto de extremo ordinario (estricto) para la función $g(x)$. Por eso, si, por ejemplo, en el punto $x^{(0)}$ la función $g(x)$ satisface las condiciones suficientes para la existencia de un extremo estricto, en este punto la función $f_0(x)$ tendrá extremo estricto condicionado respecto de las ecuaciones de enlace (43.3). Las condiciones suficientes para un extremo estricto ordinario se han obtenido anteriormente (véase el teorema 2 en el p. 40.2). Para nuestro caso tienen por expresión

1) $$\frac{\partial g(\xi^{(0)})}{\partial x_i} = 0, i = m+1, \ldots, n; \tag{43.35}$$

2) la segunda diferencial

$$d^2 g(\xi^{(0)}) = \sum_{i,j=m+1}^{n} \frac{\partial^2 g(\xi^{(0)})}{\partial x_i \partial x_j} dx_i dx_j \tag{43.36}$$

es una forma cuadrática definida positiva o definida negativa.

Si se cumplen estas condiciones, $\xi^{(0)}$ es un punto de mínimo estricto o de máximo estricto para la función $g(x)$. En virtud de lo expuesto anteriormente, las condiciones mencionadas son también suficientes para que $x^{(0)}$ sea un punto de mínimo (máximo) estricto condicionado para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3). No obstante, no son cómodas para el empleo práctico, pues requieren que se conozca la función $g(\xi)$. Por ello, partiendo de las condiciones suficientes obtenidas de un extremo estricto condicionado, que se expresan mediante la función $g(\xi)$, llegaremos a las condiciones suficientes del mismo extremo, pero expresadas esta vez sólo en términos de la función de Lagrange y las ecuaciones de enlace.

Observemos ante todo que en virtud de la condición (43.6), el sistema (43.31) es resoluble y, además, unívocamente respecto de $dx_1, \ldots, dx_m$, siendo fijadas arbitrariamente $dx_{m+1}, \ldots, dx_n$. El sistema (43.31) que expresa la igualdad a cero de las diferenciales de las funciones $f_i(x)$ en el punto $x^{(0)}$:

$$df_i = 0, i = 1, 2, \ldots, m,$$

cuando se cumplen las condiciones (43.3), se escribirá en la forma breve así:

$$df = 0, \tag{43.37}$$

donde $f = (f_1, \ldots, f_m)$.

Sea $x^{(0)}$ un punto estacionario para la función de Lagrange $F(x)$ (véase (43.13)). Esto significa que $dF(x^{(0)}) = 0$, es decir, que en este punto $\nabla f_0 + \sum_{i=1}^{m} \nabla f_i = 0$.

En el teorema 2, p. 43.4*, se ha mostrado que en este caso $x^{(0)}$ es un punto estacionario para la función $g(x)$, es decir,

$$dg(x^{(0)}) = 0. \tag{43.38}$$

Expliquemos una vez más la deducción de esta fórmula y mostremos que

$$d^2g(x^{(0)}) = d^2F(x^{(0)})|_{df=0}. \tag{43.39}$$

Esta igualdad debe entenderse como una igualdad de las funciones de $n - m$ variables $dx_{m+1}, \ldots, dx_n$. En el segundo miembro de la igualdad (43.39) las variables restantes $dx_1, \ldots, dx_m$, que figuran en las expresiones de las diferenciales escritas, se determinan a partir del sistema de ecuaciones (43.37) o bien, que es lo mismo (véanse las fórmulas (43.7)),

$$dx_k = d\varphi_k(x_1, \ldots, x_{n-m}), \quad k = 1, 2, \ldots, m.$$

Aprovechando la invariación de la forma de la primera diferencial respecto de la elección de las variables y la fórmula (43.8), tenemos

$$dg(x^{(0)}) = \sum_{j=1}^{n} \frac{\partial f_0(x^{(0)})}{\partial x_j} dx_j.$$

Sumemos a esta igualdad la suma (igual a cero) de los primeros miembros de las identidades (43.31) multiplicadas, respectivamente, por las constantes λ_i que integran la función de Lagrange $F(x)$ (con más precisión, la i-ésima igualdad (43.31) se multiplica por la constante λ_i). Luego, al hacer uso de la condición (43.13), obtendremos

$$dg(x^{(0)}) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x)\right] dx_j\Big|_{x=x^{(0)}} =$$

$$= \sum_{j=1}^{n} \frac{\partial F(x^{(0)})}{\partial x_j} dx_j = 0.$$

La afirmación (43.38) queda demostrada.

La igualdad (43.39) se demuestra del modo análogo. Escribamos, ante todo, la segunda diferencial para la función $g(x)$ en el punto $x^{(0)}$:

$$d^2g(x^{(0)}) = \sum_{j,k=1}^{n} \frac{\partial^2 f_0(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial f_0(x^{(0)})}{\partial x_j} d^2x_j. \tag{43.40}$$

Luego, al haber derivado las identidades que se obtienen como resultado de derivar las ecuaciones de enlace (43.3), es decir, las identidades

$$\frac{\partial f_i}{\partial x_1} dx_1 + \ldots + \frac{\partial f_i}{\partial x_n} dx_n = 0, \; i = 1, 2, \ldots, n,$$

tendremos en el punto $x^{(0)}$:

$$\sum_{j,k=1}^{n} \frac{\partial^2 f_i(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial f_i(x^{(0)})}{\partial x_j} d^2x_j = 0, \; i = 1, 2, \ldots, m. \tag{43.41}$$

Al multiplicar la i-ésima igualdad (43.41) por la constante λ_i, que figura en la función de Lagrange $F(x)$, sumemos las expresiones obtenidas al segundo miembro de la igualdad (43.40); en este caso tendremos

$$d^2g(\bar{x}^{(0)}) = \sum_{j,k=1}^{n} \frac{\partial^2 F(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial F(x^{(0)})}{\partial x_j} d^2x_j,$$

donde dx_i, $i = 1, 2, \ldots, n$, satisface el sistema de ecuaciones (43.37). Por cuanto el punto $x^{(0)}$ es estacionario para la función de Lagrange, el segundo miembro de la igualdad obtenida se reduce a cero y de este modo la fórmula (43.39) queda demostrada.

Diremos que la forma cuadrática $d^2F(x^{(0)})$ es cuadrática definida positiva (negativa) de las variables dx_i, $i = 1, 2, \ldots, n$, a condición de que estas variables satisfacen el sistema de ecuaciones (43.37) siempre que para cualesquiera dx_i, $i = 1, 2, \ldots, n$, que satisfacen este sistema de ecuaciones y son tales que $\sum_{i=1}^{n} (dx_i)^2 > 0$, se verifica la desigualdad $d^2F(x^{(0)}) > 0$ ($d^2F(x^{(0)}) < 0$, respectivamente).

Supongamos que el punto $x^{(0)}$ satisface las ecuaciones de enlace (43.3) y es estacionario para la función de Lagrange (43.13) y supongamos también que la segunda diferencial de la función de Lagrange en este punto es una forma cuadrática definida positiva (negativa) de las variables $dx_1, \ldots, dx_n$ a condición de que ellas satisfacen el sistema de ecuaciones (43.37). Entonces, de (43.38) y (43.39) proviene que $\bar{x}^{(0)}$ es un punto estacionario para la función $g(\bar{x})$ y que la segunda diferencial de esta función en el punto $\bar{x}^{(0)}$ es una forma cuadrática definida positiva (negativa) de las variables $dx_{m+1}, \ldots, dx_n$, y, por consiguiente, la función $g(\bar{x})$ tiene en el punto $\bar{x}^{(0)}$ un mínimo (máximo) estricto; quiere decir que la función $f_0(x)$ admite en el punto $x^{(0)}$ un mínimo (máximo) estricto condicionado respecto de las ecuaciones de enlace (43.3).

Enunciemos el resultado obtenido en forma del teorema.

Teorema 3. *Si $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ satisface las ecuaciones de enlace* (43.3) *y es un punto estacionario para la función de Lagrange* (43.13) *y si la segunda diferencial de la función de Lagrange en este punto es la forma cuadrática definida positiva (negativa) de las variables $dx_1, \ldots, dx_n$ a condición de que éstas satisfacen el sistema de ecuaciones* (43.31), *entonces $x^{(0)}$ es un punto de mínimo (máximo) estricto condicionado para la función f_0 respecto de las ecuaciones de enlace* (43.3).

De este modo, con el fin de investigar el extremo condicionado de un punto estacionario de la función de Lagrange (43.13), se debe analizar si es o no definida la forma cuadrática (43.39), es decir, examinar la segunda diferencial de la función de Lagrange en este punto, siendo cumplidas las condiciones de enlace (43.3) (cuando las diferenciales dx_i, $i = 1, 2, \ldots, n$, están ligadas entre sí por las relaciones (43.31)). En este caso hemos de tener en cuenta que si la segunda diferencial de la función de Lagrange en el punto que se considera resulta ser definida positiva (negativa) sin que se cumplan las condiciones de enlace, lo será, por supuesto, cuando dichas condiciones se cumplan.

Supongamos, por ejemplo, que se requiere hallar los puntos de extremo de la función $f(x, y) = xy$, cuando el punto (x, y) se dispone en la recta $x - y = 0$. La función de Lagrange será, en este caso, $F(x, y) = xy - \lambda(x - y)$, y, como $\frac{\partial F}{\partial x} = y - \lambda$, $\frac{\partial F}{\partial y} = x + \lambda$, entonces para la determinación de los puntos estacionarios de la función $F(x, y)$ que satisfagan las condiciones de enlace tenemos un sistema de ecuaciones

$$\begin{aligned} x - y &= 0, \\ y - \lambda &= 0, \\ x + \lambda &= 0, \end{aligned}$$

de las cuales se desprende que $x = y = \lambda = 0$.

Investiguemos en el punto $(0, 0)$ la segunda diferencial de la función $F(x, y)$, siendo cumplidas las condiciones de enlace, es decir, cuando $dx - dy = 0$. Tenemos

$$d^2F = 2dx\,dy, \tag{43.42}$$

y, por consiguiente, si se cumplen las condiciones de enlace,

$$d^2F = dx^2 \geqslant 0, \tag{43.43}$$

es decir, la segunda diferencial (43.42), representando una forma cuadrática indefinida, se convierte, siempre que se cumplan las condiciones de enlace, en una forma cuadrática definida positiva (43.43). Por eso $(0, 0)$ es el punto de mínimo estricto condicionado para el problema que acabamos de considerar. Esto, además, se ve claramente de lo siguiente: a lo largo de la recta $x - y = 0$ la función $f(x, y) = xy$ tomará la forma $f(x, x) = x^2$, admitiendo, evidentemente, en el punto $x = 0$ un mínimo estricto.

Ejercicios: Hállense los puntos de extremo condicionado de las funciones con las ecuaciones indicadas de enlace:

1. $z = \frac{x}{a} + \frac{y}{b}$, $x^2 + y^2 = 1$.
2. $z = x^2 + y^2$, $\frac{x}{a} + \frac{y}{b} = 1$.
3. $u = xyz$, $x^2 + y^2 + z^2 = 1$, $x + y + z = 0$.
4. $u = x^2 + y^2 + z^2$, $Ax + By + Cz + D = 0$.
5. $u = 2x + y - z + 1$, $x^2 + y^2 - 2z^2 - 22 = 0$.
6. Hállese el valor máximo de la función $z = xy$ en el círculo unitario $x^2 + y^2 \leqslant 1$.
7. Inscríbase en un círculo de radio dado un n-ágano de área máxima.
8. Represéntese el número $a > 0$ en forma de una suma de los sumandos $x_1, \ldots, x_n$ de un modo tal que el producto $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$ $(\alpha_i > 0,\ i = 1, \ldots, n)$ asuma el valor máximo.

CAPÍTULO SEXTO

CÁLCULO INTEGRAL DE LAS FUNCIONES DE VARIAS VARIABLES

§ 44. INTEGRALES MÚLTIPLES

44.1. CONCEPTO DE VOLUMEN EN UN ESPACIO n-DIMENSIONAL (MEDIDA DE JORDAN). CONJUNTOS MEDIBLES

Recordemos brevemente los conceptos principales relacionados con la definición de volumen n-dimensional (área, cuando $n = 2$) y demos a conocer la nueva definición de volumen (medida) de un conjunto la que se diferenciará de la introducida antes (véase el p. 31.1).

Sea R^n un espacio euclídeo n-dimensional ($n = 1, 2, 3, \ldots$). Sus puntos se designarán, como siempre, con $x = (x_1, \ldots, x_n)$, donde x_i, $i = 1, 2, \ldots, n$, son las coordenadas del punto x en cierto sistema de coordenadas fijado de una vez y para siempre. Fijemos un número entero no negativo k, ($k = 0, 1, \ldots$). Consideraremos el i-ésimo eje coordenado ($i = 1, 2, \ldots, n$), es decir, un conjunto de puntos x cuyas coordenadas son $x_1 = \ldots = x_{i-1} = x_{i+1} = \ldots = x_n = 0$. Por los puntos del eje que tienen las coordenadas del tipo $x_i = 10^{-k}m$, $m = 0, \pm 1, \pm 2, \ldots$ tracemos unos hiperplanos de dimensión $n - 1$ ortogonales al eje mencionado. El conjunto de todos los hiperplanos de esta índole construidos para todos los ejes coordenados x_i, $i = 1, 2, \ldots, n$, engendra una familia de los cubos cerrados n-dimensionales del tipo

$$Q^n = \left\{x : \frac{m_i}{10^k} \leqslant x_i \leqslant \frac{m_i + 1}{10^k}, i = 1, 2, \ldots, n\right\}, \tag{44.1}$$

donde m_i, $i = 1, 2, \ldots, n$, recorren independientemente uno del otro el conjunto de todos los números reales.

Los cubos (44.1) se llaman *cubos de rango* k, la totalidad de ellos se designa con T_k, $k = 0, 1, \ldots$.

El conjunto de todos los cubos de rango k cubre, obviamente, todo el espacio, es decir,

$$R^n = \bigcup_{Q^n \in T_k} Q^n.$$

Dos cubos de un mismo rango pueden tener en calidad de puntos comunes sólo algunos de sus puntos de frontera. Cuando $n = 1$, el cubo (44.1) es, evidentemente, un segmento y cuando $n = 2$, un cuadrado.

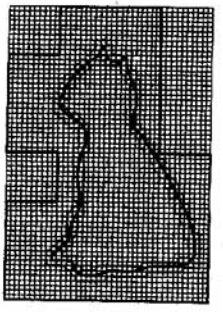


Fig. 175

El número $1/10^{kn}$ lleva el nombre de volumen n-dimensional del cubo (44.1) y se denota con μQ^n:

$$\mu Q^n \stackrel{\text{def}}{=} 10^{-kn}.$$

Para un conjunto S, que representa la unión de un número finito o numerable de diferentes cubos Q_j^n de rango dado k, $j = 1, 2, \ldots$:

$$S = \bigcup_j Q_j^n,\ Q_j^n \in T_k,$$

su volumen n-dimensional μS se determina por la igualdad

$$\mu S \stackrel{\text{def}}{=} \sum_j \mu Q_j^n. \tag{44.2}$$

Por lo visto, μS es un número no negativo o $+\infty$.

Sea ahora X un conjunto arbitrario en R^n. Designemos mediante $s_k = s_k(X)$ el conjunto de puntos de todos los cubos n-dimensionales de rango k íntegramente dispuestos en X, y mediante $S_k = S_k(X)$, el conjunto de puntos de todos los cubos n-dimensionales de rango k, cada uno de los cuales se interseca con el conjunto X por un conjunto no vacío ($k = 0, 1, 2, \ldots$):

$$s_k(X) = \bigcup_{Q^n \subset X} Q^n,\ S_k(X) = \bigcup_{Q^n \cap X \neq \varnothing} Q^n,\ Q^n \in T_k.$$

De este modo, todos los cubos de rango k, contenidos en s_k, se disponen en el conjunto X, mientras que los cubos de rango k, contenidos en S_k, forman un recubrimiento del conjunto X (fig. 175), es decir, $s_k(X) \subset X \subset S_k(X)$. Cabe notar que el conjunto X se dispone "estrictamente dentro" del poliedro $S_k = S_k(X)$, es decir, no se interseca con su frontera ∂S_k. Efectivamente, el punto $x \in X \cap \partial S_k$ no puede existir, puesto que si fuera punto de frontera para S_k, pertenecería a la cara de cierto

cubo de rango k. Ya que los cubos en consideración son cerrados, entonces, por definición del poliedro S_k, le pertenecerían a él todos los cubos de rango k que contienen dicha cara, pues ésta última contiene el punto $x \in X$. De este modo, el punto citado no sería punto de frontera para S_k.

Es evidente que

$$S_0 \subset S_1 \subset \ldots \subset S_k \subset S_{k+1} \subset \ldots, \; S_0 \supset S_1 \supset \ldots \supset S_k \supset S_{k+1} \supset \ldots,$$

y, por consiguiente, en virtud de la definición (44.2),

$$\mu S_0 \leqslant \mu S_1 \leqslant \ldots \leqslant \mu S_k \leqslant \mu S_{k+1} \ldots, \; \mu S_0 \geqslant \mu S_1 \geqslant \ldots \geqslant \mu S_k \geqslant \mu S_{k+1} \geqslant \ldots \tag{44.3}$$

De este modo, se han obtenido dos sucesiones monótonas cuyos términos son los elementos del conjunto ampliado de números reales $\bar{R}$ (véase el p. 2.5), a saber, o bien los números reales no negativos, o bien $+\infty$. Por eso, para cualquier conjunto $X \subset R^n$ siempre existen los límites finitos o infinitos

$$\lim_{k \to \infty} \mu S_k(X) \text{ y } \lim_{k \to \infty} \mu S_k(X)$$

Definición 1. *El límite finito o infinito* $\lim_{k \to \infty} \mu S_k(X)$ *lleva el nombre de medida de Jordan*[1] *n-dimensional inferior o interior del conjunto X y se designa mediante* $\mu_* X$.

$$\mu_* X \stackrel{\text{def}}{=} \lim_{k \to +\infty} \mu S_k(X), \tag{44.4}$$

mientras que el límite $\lim_{k \to +\infty} \mu S_k(X)$ *se denomina medida de Jordan n-dimensional superior o exterior del conjunto X y se designa mediante* $\mu^* X$,

$$\mu^* X \stackrel{\text{def}}{=} \lim_{k \to +\infty} \mu S_k(X). \tag{44.5}$$

Si las medidas inferior $\mu_* X$ *y superior* $\mu^* X$ *del conjunto X son finitas y coinciden, el conjunto se llama medible según Jordan. El valor general de las medidas de Jordan inferior y superior del conjunto medible X se designa con* μX *y se llama medida de Jordan n-dimensional o volumen n-dimensional del conjunto X:*

$$\mu X = \mu_* X = \mu^* X. \tag{44.6}$$

Para un conjunto vacío debe ser, por definición, $\mu \varnothing = 0$.

A veces, en lugar de μX se escribirá $\mu_n X$, con el fin de subrayar que se trata aquí de una medida del conjunto X que se considera como el subconjunto del espacio precisamente n-dimensional.

En lo que sigue, para simplificar, la medida de Jordan se llamará simplemente *medida* y el conjunto medible según Jordan, conjunto *medible*.

Por conjunto medible, como lo muestra el mismo sentido de la palabra "medible", se entiende en las matemáticas tal conjunto puntual en R^n que pueda ser medido de tal o cual manera, es decir, a este conjunto se le puede asignar, rigiéndose por ciertas reglas, un número no negativo que representa un volumen en el caso tridimensional, una área en el caso bidimensional y una longitud, en el caso unidimen-

[1] C. Jordan (1838 — 1922), matemático francés.

sional. Si la dimensión de un espacio $n \geqslant 3$, el conjunto, medible según Jordan en este espacio, se llama *cubicable* y en el caso de $n = 2$, *cuadrable*. Los términos relativos a los conjuntos cubicable y cuadrable reflejan el hecho de que la medición mencionada arriba del conjunto se realiza por medio de los cubos, o los cuadrados, respectivamente.

Por cálculos inmediatos no es difícil comprobar que si el conjunto X representa en sí una unión de un número finito de diferentes cubos n-dimensionales ($n = 1, 2, \ldots$) de rango dado, es mediable y su medida de Jordan coincide con la medida determinada por la igualdad (44.2).

Para todo conjunto X es evidente que

$$\mu s_k(X) \geqslant 0, \quad \mu S_k(X) \geqslant 0,$$

cualquiera que sea $k = 0, 1, 2, \ldots$.

Pasando al límite para $k \to \infty$, obtenemos $\mu_* X \geqslant 0$, $\mu^* X \geqslant 0$. De aquí se deduce la siguiente propiedad de la medida de Jordan.

Propiedad 1°. *Para todo conjunto medible,* $\mu X \geqslant 0$.

Observemos ahora que en virtud de las definiciones (44.4) y (44.5), para todo conjunto X queda definida una medida de Jordan finita o infinita, inferior y superior. Además, puesto que para todo $k = 0, 1, 2, \ldots$ se verifica la desigualdad $0 \leqslant \mu s_k(X) \leqslant \mu S_k(X)$, entonces, al pasar al límite para $k \to \infty$, tendremos para cualquier conjunto X

$$0 \leqslant \mu_* X \leqslant \mu^* X.$$

De aquí proviene evidentemente que si la medida superior del conjunto X es igual a cero, $\mu^* X = 0$, el conjunto X es medible y $\mu X = 0$.

Si en el conjunto X se tiene un punto interior, existe tal número k_0 que el conjunto $s_{k_0}(X)$ será no vacío; por lo tanto, $\mu s_{k_0}(X) > 0$, de donde, en virtud de (44.3), (44.4) y (44.6), se deducirá que $\mu_* X > 0$. Efectivamente, si x es un punto interior del conjunto X, existe tal $\varepsilon > 0$, que el entorno esférico $U(x, \varepsilon)$ está contenido en X. Por ello, será suficiente tomar tal rango k_0 que la longitud de la diagonal de un cubo*) de rango k_0 sea menor que ε:

$$10^{-k_0}\sqrt{n} < \varepsilon.$$

En este caso el cubo Q^n de rango k, que contiene el punto x (por lo menos un cubo de esta especie siempre existe) se dispondrá íntegramente dentro del conjunto $s_{k_0}(X)$ (fig. 176). Por esta razón

$$\mu s_{k_0}(X) \geqslant \mu Q^n > 0.$$

De lo expuesto se infiere que la *medida de Jordan inferior de cualquier conjunto abierto G es siempre positiva:* $\mu_* G > 0$.

Observemos que el volumen de un conjunto abierto, determinado anteriormente en el p. 31.1, coincide con su medida de Jordan inferior. No obstante, para construir un análogo, bastante general de la integral de Riemann en el caso de las fun-

*) La diagonal de un cubo n-dimensional con la arista de longitud a es igual a $a\sqrt{n}$.

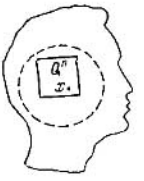


Fig. 176

ciones de varias variables, sólo el concepto de medida de Jordan inferior resulta insuficiente. Para este objeto es muy cómodo el concepto de conjunto medible según Jordan.

Si el conjunto X es acotado, $\mu_* X$ y $\mu^* X$ son siempre finitas. En efecto, del hecho de que el conjunto X es acotado proviene que dicho conjunto se interseca sólo con el conjunto finito de cubos de rango nulo y, por consiguiente, $S_0(X)$ se compone del número finito de cubos. Por ello, de acuerdo con (44.2), $\mu S_0(X) < +\infty$. Pero, para todo $k = 0, 1, \ldots$

$$s_k(X) \subset S_k(X) \subset S_0(X).$$

Por esta razón

$$0 \leqslant \mu s_k(X) \leqslant \mu S_k(X) \leqslant \mu S_0(X).$$

De aquí, pasando al límite para $k \to +\infty$, obtendremos

$$0 \leqslant \mu_* X \leqslant \mu^* X \leqslant \mu S_0(X) < +\infty,$$

es decir, las medidas $\mu_* X$ y $\mu^* X$ son finitas.

Si, en cambio, el conjunto X no es acotado, para todo $k = 0, 1, 2, \ldots$ el conjunto $S_k(X)$ se compone de un número infinito de cubos de rango k. Por eso, en virtud de la fórmula (44.2), para todo k tenemos $\mu S_k(X) = +\infty$, por consiguiente, también $\mu^* X = +\infty$, es decir, el conjunto X es a ciencia cierta no medible. De aquí:

si un conjunto es medible según Jordan quiere decir que es acotado.

Las medidas de Jordan, tanto superior como inferior, poseen la así llamada *propiedad de monotonía*. Enunciémosla en forma de un lema.

Lema 1. *Si $X_1 \subset X_2$, se tiene*

$$\mu_* X_1 \leqslant \mu_* X_2, \mu^* X_1 \leqslant \mu^* X_2. \tag{44.7}$$

Esto se deduce directamente del hecho de que para todo $k = 0, 1, 2, \ldots$ tienen lugar las inclusiones

$$s_k(X_1) \subset s_k(X_2), \quad S_k(X_1) \subset S_k(X_2), \tag{44.8}$$

pues la primera de ellas significa que un cubo de rango k dispuesto en X_1 se dispone también en X_2, mientras que la segunda es indicio de que un cubo de rango k que se interseca con el conjunto X_1 se corta también con X_2. Ambas afirmaciones pro-

vienen de la inclusión $X_1 \subset X_2$. De (44.8), en virtud de (44.2), se desprende la validez de las desigualdades

$$\mu S_k(X_1) \leqslant \mu S_k(X_2), \quad \mu S_k(X_1) \leqslant \mu S_k(X_2).$$

Al hacer tender k hacia $+\infty$, obtenemos en límite (44.7). □

Corolario 1. *Si $X_1 \subset X_2$ y $\mu X_2 = 0$, se tiene $\mu X_1 = 0$.*

Efectivamente, en vista del lema 1

$$0 \leqslant \mu^* X_1 \leqslant \mu^* X_2 = \mu X_2 = 0.$$

Por tanto, $\mu^* X_1 = 0$, de donde también $\mu X_1 = 0$. □

Corolario 2. *Si $\mu X = 0$ y $\overline{X}$ es una clausura del conjunto X* (véase el p. 18.2), *entonces $\mu \overline{X} = 0$.*

En efecto, de la condición $\mu X = 0$ se infiere que para todo $\varepsilon > 0$ existe un rango k tal que

$$\mu S_k(X) < \varepsilon.$$

El poliedro $S_k(X)$ se compone de un número finito de cubos cerrados (si el número de cubos de rango k contenidos en el conjunto $S_k(X)$ fuera infinito, entonces, de acuerdo con (44.2), su medida sería infinita: $\mu S_k(X) = +\infty$) y, por ende, es un conjunto cerrado: $\overline{S_k(X)} = S_k(X)$. Pero, $X \subset S_k(X)$, por lo cual $\overline{X} \subset \overline{S_k(X)} = S_k(X)$. De aquí $\mu^* X \leqslant \mu^* S_k(X) = \mu S_k(X)$, es decir, para cualquier $\varepsilon > 0$:

$$\mu^* \overline{X} < \varepsilon.$$

Esto es posible sólo cuando $\mu \overline{X} = 0$. □

Del lema 1 se desprende la siguiente propiedad para los conjuntos medibles.

Propiedad 2° (monotonía de la medida). *Si X_1 y X_2 son unos conjuntos medibles según Jordan y $X_1 \subset X_2$, entonces*

$$\mu X_1 \leqslant \mu X_2. \tag{44.9}$$

Lema 2 (semiaditividad de la medida superior). *Para toda totalidad finita de conjuntos $X_1, X_2, \ldots, X_m$ tiene lugar la desigualdad*

$$\mu^* \bigcup_{j=1}^{m} X_j \leqslant \bigcup_{j=1}^{m} \mu^* X. \tag{44.10}$$

DEMOSTRACIÓN. Para cualquier rango $k = 0, 1, 2, \ldots$ se verifica la igualdad

$$S_k\left(\bigcup_{j=1}^{m} X_j \right) = \bigcup_{j=1}^{m} S_k(X_j).$$

En efecto, cada cubo de rango k que se interseca con el conjunto $\bigcup_{j=1}^{m} X_j$, se interseca por lo menos con uno de los conjuntos X_j y viceversa. Por ello, en virtud de (44.2)

$$\mu S_k\left(\bigcup_{j=1}^{m} X_j \right) = \mu \bigcup_{j=1}^{m} S_k(X_j) \leqslant \sum_{j=1}^{m} \mu S_k(X_j).$$

Pasando aquí al límite para $k \to +\infty$, obtenemos (44.10). □

Corolario. *La unión de un número finito de los conjuntos de medida cero tiene medida cero.*

En efecto, si $\mu X_j = 0, j = 1, 2, \ldots, m$, entonces, en virtud de (44.10)

$$\mu^* \bigcup_{j=1}^{m} X_j \leqslant \sum_{j=1}^{m} \mu^* X_j = \sum_{j=1}^{m} \mu X_j = 0.$$

Por consiguiente, el conjunto $\bigcup_{j=1}^{m} X_j$ es medible y su medida superior y, por tanto, la medida son iguales a cero:

$$\mu \bigcup_{j=1}^{m} X_j = 0. \square$$

Ejercicios. 1. Pruébese que la unión de una totalidad numerable de los conjuntos cuya medida de Jordan es cero puede no tener la medida cero.

2. Demuéstrese que si X_1 y X_2 son los conjuntos abiertos, entonces

$$\mu_*(X_1 \cup X_2) \leqslant \mu_* X_1 + \mu_* X_2.$$

Indicación. Es útil aprovechar la afirmación contenida en el ejercicio 14 del p. 18.3.

¿Será siempre válida esta desigualdad, es decir, sin la suposición de que los conjuntos X_1 y X_2 son abiertos?

3. Dése un ejemplo de tales conjuntos disjuntos X_1 y X_2 que

$$\mu^*(X_1 \cup X_2) \neq \mu^* X_1 + \mu^* X_2.$$

El criterio de mensurabilidad de los conjuntos se establece mediante el siguiente teorema.

Teorema 1. *Para que el conjunto* X *sea medible según Jordan, es necesario y suficiente que sea acotado y que su frontera* ∂X *tenga medida de Jordan igual a cero:*

$$\mu \partial X = 0. \tag{44.11}$$

Cualquiera que sea el conjunto X, designemos mediante $\sigma_k = \sigma_k(X)$ el conjunto de puntos de aquellos, y sólo aquellos, cubos de rango k que están contenidos en $S_k(X)$ y no se contienen en $s_k(X)$:

$$\sigma_k(X) = \bigcup_{Q^n \subset S_k,\ Q^n \not\subset s_k} Q^n.$$

De este modo, el conjunto $\sigma_k(X)$ se compone de los cubos cerrados y la diferencia en el sentido de la teoría de conjuntos $S_k(X) \setminus s_k(X)$ se contiene en el conjunto $\sigma_k(X)$ y, en el caso general, no coincide con éste. Por otra parte

$$\overline{S_k(X) \setminus s_k(X)} = \sigma_k(X)^{*)}.$$

Antes de demostrar el teorema demos a conocer un lema.

---

*)La raya por encima del conjunto significa, como siempre, su clausura (véase el p. 18.2)

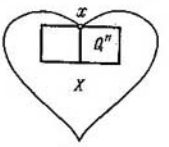


Fig. 177

Lema 3. *Para todo conjunto acotado $X \subset R^n$ son válidas las inclusiones*

$$\partial X \subset \sigma_k(X) \subset S_k(\partial X). \tag{44.12}$$

DEMOSTRACIÓN DEL LEMA. Mostremos primero que

$$\partial X \subset \sigma_k(X). \tag{44.13}$$

Por cuanto $X \subset S_k(X)$, se tiene $\overline{X} \subset \overline{S_k(X)}$. El conjunto $S_k(X)$ se compone, por ser acotado el conjunto X, de un número finito de cubos cerrados y por esta razón es cerrado: $\overline{S_k(X)} = S_k(X)$. Por consiguiente, para cualquier $k = 0, 1, 2, \ldots$, $\overline{X} \subset S_k(X)$, por ende, también $\partial X \subset S_k(X)$, pues $\partial X \subset \overline{X}$.

Tomemos un punto de frontera x del conjunto X: $x \in \partial X$. En virtud de la inclusión $\partial X \subset S_k(X)$ existe por lo menos un cubo Q^n de rango k tal que $x \in Q^n$ y $Q^n \subset S_k(X)$. Si Q^n no está contenido en $s_k(X)$, entonces, por lo visto, $Q^n \subset \sigma_k(X)$ y, por tanto, también $x \in \sigma_k(X)$.

Si, en cambio, $Q^n \subset s_k(X)$ (fig. 177), entonces, en virtud de las inclusiones $x \in Q^n$ y $Q^n \subset s_k(X) \subset X$, tenemos $x \in X$. Por eso, en este caso, todos los cubos de rango k que contienen el punto x, se disponen en $S_k(X)$, pues la intersección de todo cubo de esta índole con el conjunto X contiene el punto x y, por tanto, no es vacía. Todos estos cubos no pueden pertenecer al conjunto X, de lo contrario el punto x no sería punto de frontera del conjunto X, sino el interior. Por ello, entre todos los cubos de rango k que contienen el punto x existe por lo menos un cubo Q_0^n que no se contiene en $s_k(X)$, es decir, $Q_0^n \subset S_k(X)$, pero $Q_0^n \not\subset s_k(X)$. De aquí se deduce que $Q_0^n \subset \sigma_k(X)$, y, dado que $x \in Q_0^n$, entonces también en este caso $x \in \sigma_k(X)$. El punto x fue punto arbitrario de la frontera ∂X, y por eso, la inclusión (44.13) se ha demostrado.

La segunda inclusión (44.12), es decir, la inclusión $\sigma_k(X) \subset S_k(\partial X)$ se demuestra de un modo más fácil, sin que incluso se suponga la acotación del conjunto X. Todo cubo Q^n de rango k, que se dispone en $\sigma_k(X)$ tiene, a ciencia cierta, tanto los puntos del conjunto X (pues, en virtud de la definición de conjunto $\sigma_k(X)$, cualquier cubo de rango k, contenido en este conjunto, se contiene también en $S_k(X)$, y, por lo tanto, se interseca con X) como aquellos que no pertenecen a X (pues, de conformidad con la misma definición, ningún cubo de rango k, dispuesto íntegramente en X, es decir, perteneciente a $s_k(X)$, no se contiene en $\sigma_k(X)$). Como el cubo Q^n es un conjunto linealmente conexo, hay en él, a ciencia cierta, unos puntos de la frontera del conjunto X (véase el lema 9 en el p. 18.2). Esto significa precisamente que $Q^n \subset S_k(\partial X)$ y, como Q^n fue un cubo arbitrario de rango k, dispuesto en $\sigma_k(X)$, se tiene

$$\sigma_k(X) \subset S_k(\partial X). \square \tag{44.14}$$

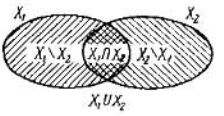


Fig. 178

DEMOSTRACIÓN DEL TEOREMA. NECESIDAD. Sea X un conjunto medible. De acuerdo con lo demostrado, es acotado. Luego, conforme a la definición de conjunto medible, las medidas inferior y superior del conjunto X son finitas e iguales: $\mu_* X = \mu^* X$, es decir,

$$\lim_{k \to +\infty} \mu s_k(X) = \lim_{k \to +\infty} \mu S_k(X). \tag{44.15}$$

Por cuanto, conforme a la definición del conjunto $\sigma_k(X)$ y la fórmula (44.2),

$$\mu\sigma_k(X) = \mu S_k(X) - \mu s_k(X), \tag{44.16}$$

de (44.15) se infiere que

$$\lim_{k \to +\infty} \mu\sigma_k(X) = 0. \tag{44.17}$$

En virtud de la inclusión (44.13) y la monotonía de la medida superior (véase (44.7), para todo $k = 0, 1, 2, \ldots$ se verifica la desigualdad

$$\mu^* \partial X \leqslant \mu^* \sigma_k(X) = \mu\sigma_k(X).$$

Pasando al límite para $k \to +\infty$, en vista de (44.17), obtendremos $\mu^* \partial X = 0$. Por lo tanto, el conjunto ∂X es medible según Jordan y $\mu\partial X = 0$.

SUFICIENCIA. Sea X un conjunto acotado y $\mu\partial X = 0$. Por definición de la medida

$$\lim_{k \to +\infty} \mu S_k(\partial X) = 0. \tag{44.18}$$

En virtud de la inclusión (44.14) y la monotonía de la medida (véase la propiedad 2 de la medida) se verifica la desigualdad $\mu\sigma_k(X) \leqslant \mu S_k(\partial X)$, y por lo tanto (véase (44.16)), la desigualdad

$$\mu S_k(X) - \mu s_k(X) \leqslant \mu S_k(\partial X). \tag{44.19}$$

Por cuanto el conjunto X es acotado, su medida inferior $\mu_* X$ y la superior $\mu^* X$ son finitas y, por ende (véanse (44.4) y (44.5)) en la desigualdad (44.19) podemos pasar al límite para $k \to +\infty$. En vista de (44.18), obtendremos

$$\mu^* X - \mu_* X = 0, \text{ es decir, } \mu^* X = \mu_* X.$$

Esto significa precisamente la mensurabilidad según Jordan del conjunto X. □

Con ayuda del teorema 1 se muestra fácilmente que al realizarse las operaciones de unión en el sentido de la teoría de conjuntos, como también las de intersección y sustracción su mensurabilidad no se perturba. Observemos previamente que para cualesquiera dos conjuntos X_1 y X_2, dispuestos en el espacio R^n, son válidas las inclusiones (fig. 178).

$$\partial(X_1 \cup X_2) \subset \partial X_1 \cup \partial X_2, \tag{44.20}$$

$$\partial(X_1 \cap X_2) \subset \partial X_1 \cup \partial X_2, \tag{44.21}$$

$$\partial(X_1 \setminus X_2) \subset \partial X_1 \cup \partial X_2, \tag{44.22}$$

Demostremos, por ejemplo, la inclusión (44.21). Sea $x \in \partial(X_1 \cap X_2)$. Entonces, ante todo, $x \in \overline{X}_1 \cap \overline{X}_2$, pues de lo que $x \in \partial(X_1 \cap X_2)$ se deduce que en todo entorno del punto x se tienen puntos que pertenecen simultáneamente a X_1 y a X_2, es decir, x es un punto de adherencia tanto del conjunto X_1, como del X_2. Si $x \in \partial X_1$ o bien $x \in \partial X_2$, o bien tienen lugar ambos casos, entonces, evidentemente, $x \in \partial X_1 \cup \partial X_2$. Si, en cambio, $x \notin \partial X_1$ y $x \notin \partial X_2$, entonces, como $x \in \overline{X}_1$ y $x \notin \partial X_1$, x es un punto interior para el conjunto X_1 y, por analogía, punto interior para el conjunto X_2 (pues la clausura de todo conjunto sólo se compone de los puntos interiores de este conjunto y sus puntos de frontera; cada uno de ellos puede ser, naturalmente, vacío). En este caso el punto x cuenta con los entornos $U_1(x) \subset X_1$ y $U_2(x) \subset X_2$, cuya intersección $U(x) = U_1(x) \cap U_2(x)$ será también el entorno del punto x, y, evidentemente, $U(x) \subset X_1 \cap X_2$. De este modo, para el punto x se ha encontrado un entorno $U(x)$, todos los puntos del cual pertenecen al conjunto $X_1 \cap X_2$, es decir, x es un punto interior y no de frontera de dicho conjunto: $x \notin \partial(X_1 \cap X_2)$. La contradicción obtenida muestra que el caso en que $x \notin \partial X_1$ y, a la vez, $x \notin \partial X_2$ es imposible, siempre que $x \in \partial(X_1 \cap X_2)$.

Ejercicio 4. Demuéstrense las inclusiones (44.20) y (44.22).

De las inclusiones (44.20) y (44.21) se establece con facilidad la validez de las inclusiones

$$\partial \bigcup_{j=1}^{m} X_j \subset \bigcup_{j=1}^{m} \partial X_j, \quad \partial \bigcap_{j=1}^{m} X_j \subset \bigcap_{j=1}^{m} \partial X_j, \tag{44.23}$$

empleando el método de inducción matemática para cualquier número finito de conjuntos.

Propiedad 3°. *La unión e intersección de un número finito de los conjuntos medibles según Jordan, como también la diferencia entre dos tales conjuntos, son conjuntos medibles según Jordan.*

Efectivamente, si los conjuntos X_j son medibles, entonces, de acuerdo con el teorema 1, $\mu\partial X_j = 0$, $j = 1, 2, \ldots, m$. Por ello, debido al corolario del lema 2, $\mu \bigcup_{j=1}^{m} \partial X_j = 0$, y en este caso (véase el corolario 1 del lema 1) de las inclusiones (44.23) se deduce, respectivamente, que

$$\mu\partial \bigcup_{j=1}^{m} X_j = 0, \quad \mu\partial \bigcap_{j=1}^{m} X_j = 0.$$

De aquí proviene que, en virtud del mismo teorema 1, los conjuntos $\bigcup_{j=1}^{m} X_j$ y

$\bigcap_{j=1}^{m} X_j$ son también medibles. Análogamente se demuestra la mensurabilidad de la diferencia entre los conjuntos medibles.

Ahora podemos demostrar fácilmente que para la medida de Jordan se verifica la desigualdad análoga a la desigualdad (44.10) para la medida superior. Enunciemos la afirmación correspondiente.

Para cualquier totalidad finita de conjuntos medibles $X_1, X_2, \ldots, X_m$ es válida la desigualdad

$$\mu \bigcup_{j=1}^{m} X_i \leqslant \sum_{i=1}^{m} \mu X_i. \tag{44.24}$$

En efecto, si los conjuntos X_i son medibles, se tiene $\mu^* X_i = \mu X_i$, y, de acuerdo con lo demostrado, la unión $\bigcup_{i=1}^{m} X_i$ es también medible y, por tanto, $\mu^* \bigcup_{j=1}^{m} X_i = \mu \bigcup_{j=1}^{m} X_i$. Por esta razón, la fórmula (44.24) en el caso que se considera coincide con la fórmula (44.10).

Propiedad 4° (aditividad de la medida). *La medida de una unión de un número finito de los conjuntos medibles según Jordan disjuntos dos a dos es igual a la suma de las medidas de estos conjuntos.*

De este modo, si X_i son conjuntos medibles, $X_i \cap X_j = \varnothing, i \neq j, i = 1, 2, \ldots, m$, entonces

$$\mu \bigcup_{i=1}^{m} X_i = \sum_{i=1}^{m} \mu X_i. \tag{44.25}$$

Demostremos esto. Como para cualquier rango k es válida la inclusión $s_k(X_i) \cap \cap s_k(X_j) \subset X_i \cap X_j$, de la condición $X_i \cap X_j = \varnothing$, para $i \neq j$, se infiere que $s_k(X_i) \cap \cap s_k(X_j) = \varnothing, i \neq j$; por ello, de acuerdo con (44.2),

$$\sum_{i=1}^{m} \mu s_k(X_i) = \mu \bigcup_{i=1}^{m} s_k(X_i). \tag{44.26}$$

Si el cubo de rango k se dispone en cierto conjunto X_i, se dispone también en la unión $\bigcup_{i=1}^{m} X_i$, y, por consiguiente,

$$\bigcup_{i=1}^{m} s_k(X_i) \subset s_k\left(\bigcup_{i=1}^{m} X_i\right).$$

De aquí, en virtud de (44.26) y la monotonía de la medida (en el caso dado incluso de la fórmula (44.2)) se desprende que

$$\sum_{i=1}^{m} \mu s_k(X_i) = \mu \bigcup_{i=1}^{m} s_k(X_i) \leqslant \mu s_k\left(\bigcup_{i=1}^{m} X_i\right).$$

Pasando al límite para $k \to +\infty$, obtenemos

$$\sum_{i=1}^{m} \mu X_i \leqslant \mu \bigcup_{i=1}^{m} X_i. \tag{44.27}$$

Por otra parte, para cualesquiera conjuntos medibles se verifica la desigualdad (44.24). Evidentemente, de (44.24) y (44.27) proviene la igualdad (44.25), es decir, aditividad de la medida.

OBSERVACIÓN. De las propiedades 3 y 4 se infiere que si a un conjunto medible se le suma o se le resta un conjunto de medida cero, el conjunto obtenido será también medible y la medida de éste será igual a la del conjunto de partida. Efectivamente, si X es un conjunto medible y $\mu X_0 = 0$, entonces, según la propiedad 3 de la medida, los conjuntos $X \setminus X_0$ y $X \cup X_0$ son también medibles. Luego, de conformidad con la propiedad 4, para $X_0 \subset X$ y $\mu X_0 = 0$ tenemos

$$\mu X = \mu[(X \setminus X_0) \cup X_0] = \mu(X \setminus X_0) + \mu X_0 = \mu(X \setminus X_0).$$

Tomando en consideración la monotonía de la medida y las desigualdades (44.24) llegamos a que para todo X_0, $\mu X_0 = 0$ se verifican las desigualdades

$$\mu X \leqslant \mu(X \cup X_0) \leqslant \mu X + \mu X_0 = \mu X,$$

de donde $\mu(X \cup X_0) = \mu X$.

De lo expuesto proviene, además, que si a un conjunto medible se le agrega o se le resta algún conjunto de sus puntos de frontera, se obtendrá nuevamente un conjunto medible de la misma medida que el conjunto dado. Esto se debe a que en virtud del teorema 1, la frontera de un conjunto medible y, por ende, cualquier subconjunto de ésta, tienen medida cero. De este modo, en particular, si el conjunto X es medible, su clausura $\bar{X} = X \cup \partial X$ es también medible, con la particularidad de que $\mu X = \mu \bar{X}$.

La afirmación recíproca no es cierta: *existen unos conjuntos no medibles según Jordan cuyas clausuras son medibles.* De ejemplo sencillo del conjunto semejante sirve el conjunto de todos los puntos reales en cierto segmento. Dicho conjunto es no medible (¿por qué?), mientras que su clausura está constituida por un segmento que es medible.

Los ejemplos de conjuntos medibles de dimensión tan grande como se quiera se pueden obtener construyendo cilindros cuyas bases son los conjuntos medibles. Enunciemos la definición del cilindro.

Definición 2. *Sea X_0 un conjunto dispuesto en la hipersuperficie $R^{n-1} = \{x : x_n = 0\}$ del espacio R^n, y sean a y b unos números reales, $a \leqslant b$. El conjunto*

$$X = \{x : (x_1, x_2, \ldots, x_{n-1}, 0) \in X_0,\ a \leqslant x_n \leqslant b\}$$

se llama cilindro n-dimensional de base X_0 y generatriz (paralela al eje coordenado x_n) de longitud $h = b - a$.

Es obvio que recurriendo a la noción de producto de los conjuntos (véase el p. 1.2° ó 41.2), podemos decir que el cilindro X es un producto de los conjuntos X_0 y del segmento $[a, b]$: $X = X_0 \times [a, b]$. Si X_0 es un conjunto acotado, el cilindro de

base X_0 es un conjunto acotado. De aquí se infiere que todo cilindro, de cuya base sirve un conjunto medible, es acotado, pues el conjunto medible es acotado.

Teorema 2. *Si X_0 es un conjunto medible según Jordan del espacio R^{n-1}, todo cilindro n-dimensional X de base X_0 es un conjunto medible según Jordan del espacio R^n y*

$$\mu_n X = h\mu_{n-1}X_0, \tag{44.28}$$

donde h es la longitud de la generatriz del cilindro X.

Corolario. *Si la base del cilindro tiene medida (n − 1)-dimensional igual a cero, entonces el propio cilindro n-dimensional tiene medida n-dimensional que es también igual a cero.*

DEMOSTRACIÓN DEL TEOREMA. Observemos, ante todo, que la proyección*) de cualquier cubo n-dimensional Q^n de rango k representa un cubo $(n-1)$-dimensional Q^{n-1} también de rango k y

$$\mu Q^n = 10^{-k}\mu Q^{n-1}. \tag{44.29}$$

Designemos mediante $q_1^{n-1}, \ldots, q_l^{n-1}$ los cubos $(n-1)$-dimensionales de rango k que componen el conjunto $s_k(X_0)$ y mediante $Q_1^{n-1}, \ldots, Q_m^{n-1}$, los cubos $(n-1)$-dimensionales que componen el conjunto $S_k(X_0)$.

Sean $q_{j1}^n, \ldots, q_{jp}^n$ los cubos n-dimensionales de $s_k(X)$ que se proyectan en el cubo $q_j^{n-1} \subset s_k(X_0)$. Por cuanto X es un cilindro, el número p de estos cubos n-dimensionales q_{ij}^n será el mismo para todo $i = 1, 2, \ldots, l$, por lo cual

$$s_k(X) = \bigcup_{i=1}^{l} \bigcup_{j=1}^{p} q_{ij}^n. \tag{44.30}$$

Análogamente, el número r de cubos n-dimensionales Q_{ij}^n de $S_k(X)$ que se proyectan en un mismo cubo Q_i^{n-1} de $S_k(X_0)$ es igual para todo $i = 1, 2, \ldots, m$, por lo cual

$$S_k(X) = \bigcup_{i=1}^{m} \bigcup_{j=1}^{r} Q_{ij}^n. \tag{44.31}$$

La proyección del conjunto $\bigcup_{j=1}^{p} q_{ij}^n$ sobre el eje x_n es un segmento de longitud $p \cdot 10^{-k}$, con la particularidad de que

$$p \cdot 10^{-k} \leqslant h, \tag{44.32}$$

pues todos los cubos q_{ij}^n se contienen en $s_k(X)$ y, por consiguiente, en el cilindro X. En cuanto a la proyección del conjunto mencionado sobre el hiperplano R^{n-1}, ésta representa uno de los cubos q_i^{n-1}, por lo cual

---

*) Se llama proyección $\mathrm{pr}_{x_n} X$ del conjunto $X \subset R^n$ sobre el hiperplano $R^{n-1} = \{x : x_n = 0\}$ un conjunto de puntos del tipo $(x_1, \ldots, x_{n-1}, 0)$ para cada uno de los cuales existe tal x_n que $(x_1, \ldots, x_{n-1}, x_n) \in X$.

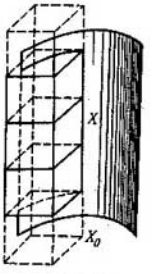


Fig. 179

$$\mu_n s_k(X) = \sum_{i=1}^{l} \sum_{j=1}^{p} \mu_n Q_{ij}^n = \sum_{i=1}^{l} \sum_{j=1}^{p} \frac{1}{10^k} \mu_{n-1} q_i^{n-1} =$$

$$= \frac{p}{10^k} \sum_{i=1}^{l} \mu_{n-1} q_i^{n-1} = \frac{p}{10^k} \mu_{n-1} s_k(X_0). \quad (44.33)$$

La proyección de la "columna de cubos" $\bigcup_{j=1}^{r} Q_{ij}^n$ (fig. 179) sobre el eje x_n es un segmento de longitud $10^{-k}r$, con la particularidad de que

$$h \leqslant \frac{r}{10^k} \leqslant h + \frac{2}{10^k}. \quad (44.34)$$

Luego, cada columna de este género se proyecta sobre el hiperplano R^{n-1} en el cubo Q_i^{n-1} que o bien se contiene en $s_k(X_0)$ o bien en $\sigma_k(X_0) = \overline{S_k(X_0)} \setminus s_k(X_0)$ (el conjunto $\sigma_k(X_0)$ se ha introducido al demostrar el teorema 1). Por esta razón

$$\mu_n S_k(X) = \sum_{i=1}^{m} \sum_{j=1}^{r} \mu_n Q_{ij}^n = \sum_{i=1}^{m} \sum_{j=1}^{r} \frac{1}{10^k} \mu_{n-1} Q_i^{n-1} =$$

$$= \frac{r}{10^k} \sum_{i=1}^{m} \mu_{n-1} Q_i^{n-1} = \frac{r}{10^k} [\mu s_k(X_0) + \mu \sigma_k(X_0)]. \quad (44.35)$$

Observemos, por fin, que cada una de las columnas $\bigcup_{j=1}^{r} Q_{ij}^n$ que se proyecta en el cubo $q^{n-1} \subset s_k(X_0)$ se diferencia de la columna $\bigcup_{j=1}^{p} q_{ij}^n$, que se proyecta en el mismo cubo, sólo en dos cubos adicionados a la columna por arriba y por de-

bajo (en el sentido de decrecimiento, de crecimiento, respectivamente, de la coordenada x_n, véase la fig. 179). Por ello

$$r = p + 2.$$

De aquí y de las desigualdades (44.32) y (44.34) tenemos

$$\frac{p}{10^k} \leqslant h \leqslant \frac{p+2}{10^k} = \frac{r}{10^k} \leqslant h + \frac{2}{10^k}.$$

Debido a esta circunstancia, de las desigualdades (44.33) y (44.35), obtenemos

$$\mu_n S_k(X) - \mu_n s_k(X) = \frac{r-p}{10^k}\mu_{n-1} s_k(X_0) + \\ + \frac{r}{10^k}\mu\sigma_k(X_0) \leqslant \frac{2}{10^k}\mu_{n-1} X_0 + \left(h + \frac{2}{10^k}\right)\mu\sigma_k(X_0)$$

y como $\lim\limits_{k\to\infty} \frac{2}{10^k} = 0$, $\lim\limits_{k\to\infty} \mu\sigma_k(X_0) = 0$, entonces

$$\lim_{k\to+\infty} [\mu_n S_k(X) - \mu_n s_k(X)] = 0. \tag{44.36}$$

El conjunto X_0 es acotado, como cualquier otro conjunto medible. No es difícil convencerse de que los diámetros $d(X_0)$ y $d(X)$ de los conjuntos X_0 y X están ligados mediante la correlación $d(X) = \sqrt{[d(X_0)]^2 + h^2}$, de la cual se deduce que el conjunto X es también acotado. Por ello, de acuerdo con lo expuesto más arriba, tiene medidas finitas superior e inferior. De las fórmulas (44.4), (44.5) y (44.36) proviene que dichas medidas son iguales: $\mu^* X = \mu_* X$, es decir, el conjunto X es medible.

Demostremos ahora la fórmula (44.28). Con este objeto multipliquemos por h la desigualdad

$$\mu_{n-1} s_k(X_0) \leqslant \mu X_0 \leqslant \mu_{n-1} S_k(X_0).$$

Al aplicar las desigualdades (44.32) y (44.34), tendremos (véase también (44.33) y (44.35))

$$\mu_n s_k(X) = \frac{p}{10^k}\mu_{n-1} s_k(X_0) \leqslant h\mu X_0 \leqslant \frac{r}{10^k}\mu_{n-1} S_k(X_0) = \mu_n S_k(X),$$

observando que en virtud de (44.36) ambos miembros de la desigualdad obtenida tienden, para $k \to \infty$, hacia un mismo límite μX, de lo que se infiere la fórmula (44.28).

Problema 29. Constrúyase un ejemplo de región que sea no medible según Jordan.

Problema 30. Demuéstrese que la medida de Jordan no depende de cómo se elige el sistema de coordenadas cartesianas.

44.2. CONJUNTOS DE MEDIDA CERO

En el punto antecedente se ha establecido que un conjunto es medible según Jordan cuando, y sólo cuando, su frontera tiene medida cero. Resulta importante, pues, conocer los criterios, rigiéndose por los cuales se podría establecer que un con-

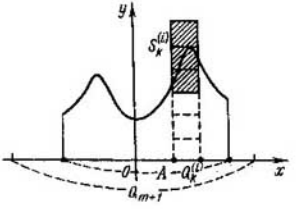


Fig. 180

junto tiene medida cero. Como ejemplo suficientemente general de conjuntos de medida cero sirven los cilindros en cuya base se disponen conjuntos de medida cero (véase el corolario del teorema 2). Otra clase amplia de los conjuntos de medida cero se proporciona en el teorema que sigue.

Teorema 3. *La gráfica de cualquier función continua en un compacto tiene medida cero.*

DEMOSTRACIÓN. Supongamos que la función $y = f(x) = f(x_1, \ldots, x_n)$ es continua en el compacto $A \subset R_x^n$. Sea X la gráfica de esta función, es decir, el conjunto de tales puntos $(x, y) = (x_1, \ldots, x_n, y)$ en el espacio n-dimensional R_{xy}^{n+1} que $(x_1, \ldots, x_n) \in A$ e $y = f(x_1, \ldots, x_n)$:

$$X = \{(x, y) : (x_1, \ldots, x_n) \in A,\ y = f(x_1, \ldots, x_n)\}.$$

Mostremos que la medida de Jordan (n + 1)-dimensional del conjunto X es igual a cero. El conjunto A, siendo un compacto, es acotado. Por eso existe tal número natural m que el cubo n-dimensional

$$P_m = \{x : -m \leqslant x_i \leqslant m,\ i = 1, 2, \ldots, n\}$$

contiene el conjunto $A : P_m \supset A$. Con mayor razón el cubo

$$P_{m+1} = \{x : -m - 1 \leqslant x_i \leqslant m + 1,\ i = 1, 2, \ldots, n\}$$

también contiene $A : P_{m+1} \supset A$, y más aún, cualquiera que sea el cubo Q de cierto rango $k = 0, 1, 2, \ldots$, que se interseca con el conjunto A, es decir, $Q \subset S_k(A)$, está contenido también en $P_{m+1} : Q \subset P_{m+1}$. Por lo tanto, para todo k se tiene $S_k(A) \subset \subset P_{m+1}$. Aquí y en adelante mediante $S_k(A)$, $S_k(X)$, al igual que en el p. 44.1, se denotan los conjuntos de puntos de todos los cubos de rango k de los espacios correspondientes que se intersecan con los conjuntos $A \subset X_x^n$, $X \subset R_{xy}^{n+1}$.

El conjunto $S_k(X)$ se descompone en un número finito de "columnas" $S_k^{(i)}$, $i = 1, 2, \ldots, i_0$, cada una de las cuales consta de los cubos $(n + 1)$-dimensionales de rango k que tienen una misma proyección (véase la llamada en la pág. 128) Q_k^i en el espacio R_x^n (en la fig. 180 se muestra el caso en que $n = 1$):

$$S_k(X) = \bigcup S_k^{(i)},\ \operatorname{pr}_y S_k^{(i)} = Q_k^{(i)},\ k = 0, 1, 2, \ldots. \tag{44.37}$$

Designemos con $\omega(\delta)$ el módulo de continuidad de la función f en A. Observando que la diagonal (el diámetro*) del cubo n-dimensional con la arista de longitud $1 \cdot 10^{-k}$ es igual a $\sqrt{n} \cdot 10^{-k}$, para la altura $h_k^{(i)}$ de cada columna $S_k^{(i)}$ tenemos (véase la fig. 161) una estimación

$$h_k^{(i)} \leqslant \omega\left(\frac{\sqrt{n}}{10^k}\right) + \frac{2}{10^k}. \tag{44.38}$$

Efectivamente, para estimar la altura $h_k^{(i)}$, a la distancia $\omega(10^{-k}\sqrt{n})$ entre los valores máximo y mínimo de la función $f(x)$ en el cubo $Q_k^{(i)}$ es suficiente añadir las longitudes de las aristas de los cubos más inferior y más superior de la columna en consideración $S_k^{(i)}$ (esta estimación se logra cuando los puntos de la gráfica, correspondientes a los valores extremos mencionados, se dispongan en las aristas, paralelas al espacio $R_{x'}^n$, de los cubos de rango k). De (44.37) y (44.38) obtenemos:

$$\mu S^k(X) = \mu \bigcup_i S_k^{(i)} = \sum_i \mu S_k^{(i)} =$$

$$= \sum_i h_k^{(i)} \mu Q_k^{(i)} \leqslant \left[\omega\left(\frac{\sqrt{n}}{10^k}\right) + \frac{2}{10^k}\right] \sum_i \mu Q_k^{(i)} \leqslant$$

$$\leqslant \left[\omega\left(\frac{\sqrt{n}}{10^k}\right) + \frac{2}{10^k}\right] \mu P_{m+1}. \tag{44.39}$$

Siendo la función f continua en el compacto, es uniformemente continua en él, por lo cual $\lim\limits_{k \to +\infty} \omega(10^{-k}\sqrt{n}) = 0$, y como $\lim\limits_{k \to +\infty} \frac{2}{10^k} = 0$, de (44.39) tenemos $\lim\limits_{k \to \infty} \mu S_k(X) = 0$, lo que es testimonio de que $\mu^* X = 0$, por consiguiente, también $\mu X = 0$. □

En virtud de los teoremas 2 y 3, todo conjunto acotado cuya frontera puede ser representada como una unión de un número finito de los conjuntos, cada uno de los cuales representa o bien una parte de la gráfica de una función continua en el conjunto cerrado acotado, o bien una parte de cilindro con la base de medida cero, es un conjunto medible, pues debido a la aditividad de la medida, la medida de la frontera del conjunto citado es igual a cero y, por lo tanto, de acuerdo con el teorema 1, el conjunto es medible. De este modo hemos obtenido la descripción de una clase bastante amplia de conjuntos medibles según Jordan que se encuentran frecuentemente en el análisis matemático y en sus aplicaciones. Así, por ejemplo, los conjuntos planos (trapecios curvilíneos, "sectores" de las curvas definidas en las coordenadas polares, como también cuerpos de revolución, cuyas áreas y, respectivamente, volúmenes, se calculaban en el § 32 con ayuda de la integral unidimensional de Riemann) son conjuntos medibles según Jordan, pues sus fronteras tienen medida cero, de lo que no es difícil convencerse.

De modo semejante son medibles según Jordan los paralelepípedos y elipsoides, en particular, las esferas, puesto que sus fronteras pueden representarse en forma de la unión de las gráficas de las funciones continuas en los compactos.

*) Véase en el p. 19.6 la definición de diámetro de un conjunto.

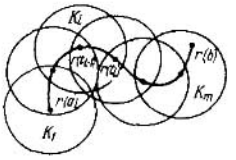


Fig. 181

Cabe notar que en el § 31 se ha introducido el concepto de medida mes G para los conjuntos abiertos. Al comparar la definición para la medida mencionada con la aducida en el p. 44.1, vemos que mes $G = \mu_* G$ es decir, la medida introducida en el § 31 es una medida inferior de Jordan. Sin embargo, en virtud de lo expuesto más arriba, todos los conjuntos considerados en los ejemplos del § 32 eran medibles según Jordan y, por consiguiente, para ellos mes G fue la medida de Jordan, es decir, para los conjuntos mencionados tuvo lugar mes $G = \mu G$.

Resulta interesante generalizar el teorema 3 para el caso de los conjuntos dados paramétricamente, en particular, para unas curvas paramétricas. Se pone de manifiesto que incluso en este caso, para que las curvas en consideración tengan medida cero no es suficiente que sean sólo continuas. Existen, por ejemplo, unas curvas $x_i = x_i(t)$, $a \leqslant t \leqslant b$, $i = 1, 2, \ldots, n$ ($x_i(t)$ son las funciones continuas en cierto segmento $[a, b]$) llamadas curvas de Peano*) que pasan por cada punto de cierto cubo n-dimensional y, por ende, no tienen medida cero.

Problema 31. Constrúyase un ejemplo de curva de Peano.

Teorema 4. *Toda curva plana rectificable tiene medida cero.*

DEMOSTRACIÓN. Sea dada una curva rectificable γ, cuya longitud es igual a S. Supongamos, además, que $r = r(t)$, $a \leqslant t \leqslant b$, es cierta representación de la curva γ. Dividamos la curva mencionada de manera sucesiva, o sea, en el orden de crecimiento del parámetro t, mediante los puntos $r(t_i)$, $t_i \in [a, b]$, $i = 0, 1, \ldots, m$, $t_0 = a$, $t_m = b$, en m partes de igual longitud, es decir, adoptemos tal partición $\tau = \{t_i\}_{i=0}^{m}$ del segmento $[a, b]$, que la longitud de cada parte γ_i (de la curva γ), definida por la representación $r = r(t)$, $t_{i-1} \leqslant t \leqslant t_i$, $i = 1, 2, \ldots, m$, tenga la longitud S/m.

Designemos mediante K_i un círculo cerrado de radio S/m y centro en el punto $r(t_{i-1})$. Ya que el arco γ_i es de longitud S/m y su origen coincide con el centro del círculo K_i, se dispone íntegramente en este círculo (fig. 181). De aquí se deduce que toda la curva γ está contenida en la reunión de los círculos K_i:

$$\gamma \subset \bigcup_{i=1}^{m} K_i.$$

*) J. G. Peano (1858 — 1932), matemático italiano.

Por consiguiente, por ser monótona y semiaditiva la medida superior (véase los lemas 1 y 2 en el p. 44.1),

$$\mu^*\gamma \leqslant \mu^* \bigcup_{i=1}^{m} K_i \leqslant \sum_{i=1}^{m} \mu^* K_i. \tag{44.40}$$

Pero, $\mu^* K_i = \mu K_i = \pi\left(\frac{S}{m}\right)^2$, $i = 1, \ldots, m$*), por lo cual, de (44.40) tenemos

$$\mu^*\gamma \leqslant \pi S^2/m.$$

El primer miembro de la desigualdad no depende de m, mientras que el segundo miembro tiende a cero cuando $m \to +\infty$, a consecuencia de lo cual $\mu\gamma = 0$. □

Ejercicio 5. Demuéstrese que toda curva rectificable tiene en un espacio tridimensional una medida cero.

De los teoremas 1 y 4 proviene que todo conjunto plano acotado, cuya frontera es una curva rectificable, es medible.

44.3. DEFINICIÓN DE LA INTEGRAL MÚLTIPLE

Enunciemos la definición de la integral múltiple de Riemann. Con este fin introduzcamos ante todo el concepto de partición de un conjunto medible y el de finura de dicha partición.

Sea X un conjunto medible según Jordan, $X \subset R^n$. Un sistema finito $\tau = \{X_i\}_{i=1}^{i_0}$ de conjuntos no vacíos medibles según Jordan X_i, $i = 1, 2, \ldots, i_0$, se denomina *partición del conjunto* X, siempre que

1) las intersecciones de dos en dos de los conjuntos X_i tienen medida cero:

$$\mu(X_i \cap X_j) = 0, \quad i \neq j;$$

2) $\bigcup_{i=1}^{i_0} X_i = X$.

El número $\delta_\tau = \max_{i=1, 2, \ldots, i_0} d(X_i)$, donde $d(X_i)$ es el diámetro del conjunto X_i, lleva el nombre de *finura de la partición* τ.

Por ser aditiva la medida de Jordan, para toda partición $\tau = \{X_i\}_{i=1}^{i_0}$ del conjunto X tenemos

$$\mu X = \sum_{i=1}^{i_0} \mu X_i. \tag{44.41}$$

En efecto, supongamos que para i fijo se tiene $X_i^* = \bigcup_{j \neq i} X_i \cap X_j$

*) En efecto, la circunferencia C, que es la frontera del círculo K, puede representarse como reunión de dos semicircunferencias, cada una de las cuales representa la gráfica de una función continua en el segmento. Por ello, de acuerdo al teorema 3, $\mu C = 0$, por consiguiente, todo círculo K es un conjunto medible.

y $X^* = \bigcup_{i=1}^{i_0} X_i^*$. En tal caso, en virtud del p. 1) de la definición de partición de un conjunto, $\mu(X_i \cap X_j) = 0, i \neq j$, por lo cual $\mu X_i^* \leqslant \sum_{j \neq 1} \mu(X_i \cap X_j) = 0$, es decir, $\mu X_i^* = 0$. De aquí $\mu X^* \leqslant \sum_{i=1}^{i_0} \mu X_i^* = 0$, por consiguiente, $\mu X^* = 0$. Además, los conjuntos $X^*, X_i \setminus X^* = X_i^{**}, i = 1, 2, \ldots, i_0$, son disjuntos dos a dos y, en virtud del p. 2), $\bigcup_{i=1}^{i_0} X_i^{**} \cup X^* = \bigcup_{i=1}^{i_0} X_i = X$.

Puesto que

$$\mu X_i = \mu(X_i \setminus X^*) = \mu X_i^{**},$$

de todo lo dicho proviene, por ser aditiva la medida, que

$$\mu X = \sum_{i=1}^{i_0} \mu X_i^{**} + \mu X^* = \sum_{i=1}^{i_0} \mu X_i. \square$$

Para simplificar las designaciones, a veces, en lugar de $\{X_i\}_{i=1}^{i_0}$ escribiremos $\{X_i\}$.

Sean $\tau = \{X_i\}$ y $\tau' = \{X_j'\}$ las particiones del conjunto medible X. La partición τ' se llama *inscrita* en la partición τ, si para cualquier $X_j' \in \tau'$ existe tal elemento de $X_i \in \tau$ que $X_j' \subset X_i$. En este caso suele escribirse $\tau' > \tau$, o bien $\tau < \tau'$.

Cabe señalar dos propiedades de las particiones de un conjunto.

1°. *Si* $\tau < \tau'$ *y* $\tau' < \tau''$, *se tiene* $\tau < \tau''$.

2°. *Para cualesquiera dos particiones* $\tau' = \{X_i'\}$ *y* $\tau'' = \{X_j''\}$ *de un conjunto medible* X *existe tal partición suya* τ *que* $\tau > \tau'$ *y* $\tau > \tau''$.

La propiedad 1° proviene, evidentemente, de la definición de partición inscrita. A título de partición τ, indicada en la propiedad 2°, puede tomarse un conjunto de toda una serie de intersecciones no vacías $X_i' \cap X_j''$.

Como ejemplo de partición de un conjunto medible interviene una totalidad de toda clase de intersecciones no vacías del conjunto dado con los cubos de cierto rango fijado k. De aquí se ve que para todo conjunto medible existen particiones cuya finura es tan menuda como se quiera.

Definición 3. *Sea* $y = f(x) = f(x_1, \ldots, x_n)$ *una función dada en el conjunto medible según Jordan* $X \subset R^n$ *y sea* τ *una partición del conjunto* X, $\tau = \{X_i\}_{i=1}^{i_0}$; *elijamos arbitrariamente los puntos* $\xi^{(i)} \in X_i$, $i = 1, 2, \ldots, i_0$. *Una suma del tipo*

$$\sigma_\tau = \sigma_\tau(f; \xi^{(1)}, \ldots, \xi^{(i_0)}) = \sum_{i=1}^{i_0} f(\xi^{(i)}) \mu X_i \qquad (44.42)$$

lleva el nombre de suma integral de Riemann de la función f.

Al igual que en el caso de la función de una sola variable, la definición de la integral múltiple puede enunciarse, haciendo uso del concepto de límite de una sucesión o "el lenguaje $\varepsilon - \delta$".

Definición 4. *El número A se denomina integral de Riemann de la función f referida al conjunto medible según Jordan $X \subset R^n$, si cualquiera que sea la sucesión de particiones $\tau_m = \{X_i^m\}_{i=1}^{i(m)}$, $m = 1, 2, \ldots$, del conjunto X tal que las finuras de las particiones τ_m tienden a cero cuando $m \to +\infty$: $\lim_{m \to +\infty} \delta_{\tau_m} = 0$, y cualesquiera que sean los puntos $\xi^{(i, m)} \in X_i^m$, $i = 1, 2, \ldots, i^{(m)}$, la sucesión de sumas integrales $\sigma_{\tau_m}(f; \xi^{(1, m)}, \ldots, \xi^{(i^{(m)}, m)})$ tiene como su límite, para $m \to +\infty$, el número A:*

$$\lim_{m \to +\infty} \sigma_{\tau_m}(f; \xi^{(1, m)}, \ldots, \xi^{(i^{(m)}, m)}) = A. \tag{44.43}$$

La integral de la función f, referida al conjunto X, se designa mediante

$$\int f(x)dX \text{ o bien } \int\int \ldots \int_X f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n.$$

Si existe la integral $\int f(x)dX$, la función f se llama *integrable según Riemann* en el conjunto X. Las funciones integrables según Riemann se denominarán con frecuencia *integrables*.

La igualdad (44.43), es decir, la definición de la integral, se escribe brevemente como la fórmula

$$\int f(x)dX = \lim_{\sigma_\tau \to 0} \sigma_\tau. \tag{44.44}$$

En términos de ε y δ este límite significa lo siguiente: *para cualquier $\varepsilon > 0$ existe tal $\delta_\varepsilon > 0$, que cualquiera que sea la partición $\tau = \{X_i\}_{i=1}^{i_0}$ del conjunto X de finura $\delta_\tau < \delta_\varepsilon$ y cualesquiera que sean los puntos $\xi^{(i)} \in X_i$, $i = 1, 2, \ldots, i_0$, se verifica la desigualdad*

$$|\sigma_\tau(f; \xi^{(1)}, \ldots, \xi^{(i_0)}) - \int f(x)dX| < \varepsilon. \tag{44.45}$$

Recurriendo al modo habitual, se demuestra que las definiciones (44.43) y (44.45) del límite de sumas integrales son equivalentes.

Señalemos que la definición de la integral (44.44) en el caso en que $n = 1$ y a título de conjunto, respecto al cual se realiza la integración, sirve un segmento, no coincide formalmente con la definición de integral de Riemann, introducida anteriormente, de la función de una sola variable, puesto que en aquel caso se consideraban sólo las particiones de un segmento en segmentos, mientras que ahora se consideran toda clase de particiones de un segmento en conjuntos medibles según Jórdan. No obstante, se puede mostrar (lo haremos en el p. 44.7*) que para $n = 1$ ambas definiciones resultan equivalentes al considerar el caso en que un conjunto, respecto del cual se realiza la integración, es un segmento, es decir, las dos definiciones conducen a un mismo concepto de integrabilidad de las funciones y a un mismo concepto de integral.

Al determinar una integral referida al conjunto $X \subset R^n$, se pueden utilizar, para formar las sumas integrales, no todos los elementos de las particiones τ del conjunto X, sino despreciar aquellos sumandos que corresponden a los elementos de la partición cuyas clausuras se intersecan con cierto conjunto fijo de medida cero. Analicemos esta circunstancia más detalladamente.

Sea X un conjunto medible, $X_0 \subset X$, y sea $\tau = \{X_i\}_{i=1}^{i=p}$ la partición del conjunto X. Designemos mediante $\tau(X_0)$ la totalidad de aquellos elementos de la partición τ, cuyas clausuras no se intersecan con el conjunto X_0:

$$\tau(X_0) = \{X_i : \overline{X}_i \cap X_0 = \varnothing, X_i \in \tau\} \tag{44.46}$$

y, viceversa, con $\tau_0(X_0)$, la totalidad de aquellos X_i, para los cuales sus clausuras $\overline{X}_i$ se cortan con X_0:

$$\tau_0(X_0) = \{X_i : \overline{X}_i \cap X_0 \neq \varnothing, X_i \in \tau\}. \tag{44.47}$$

Lema 6. *Sea X un conjunto del espacio R^n medible según Jordan, $X_0 \subset X$ y $\mu X_0 = 0$. En este caso*

$$\lim_{\delta_\tau \to 0} \sum_{X_i \in \tau_0(X_0)} \mu X_i = 0. \tag{44.48}$$

Esta igualdad significa que para todo $\varepsilon > 0$ existe tal $\delta > 0$ que cualesquiera que sean las particiones τ del conjunto X de finura $\delta_\tau < \delta$, se verifica la desigualdad

$$\sum_{X_i \in \tau_0(X_0)} \mu X_i < \varepsilon.$$

La sumación en la fórmula (44.48) se efectúa sólo según aquellos índices i, para los cuales $X_i \in \tau_0(X_0)$.

DEMOSTRACIÓN. Sea $X_0 \subset X$ y $\mu X_0 = 0$; entonces también $\mu \overline{X}_0 = 0$ (véase en el p. 44.1 la observación que sigue tras la demostración de la aditividad de una medida). Por ello, para todo $\varepsilon > 0$ existe tal rango k que

$$\mu S_k(\overline{X}_0) < \varepsilon. \tag{44.49}$$

Aquí, como siempre, $S_k(\overline{X}_0)$ denota la totalidad de los puntos de todos los cubos de rango k que se intersecan con el conjunto $\overline{X}_0$ y, por lo tanto, lo recubren: $\overline{X}_0 \subset \subset S_k(\overline{X}_0)$.

Recordemos que $\overline{X}_0$ se dispone estrictamente dentro del poliedro $S_k(\overline{X}_0)$, es decir, no se corta con su frontera (véase el p. 44.1). Por cuando el conjunto X_0 es acotado y cerrado, mientras que la frontera $\partial S_k(\overline{X}_0)$ del poliedro $S_k(\overline{X}_0)$, igual que la frontera de cualquier conjunto, está cerrada, entonces $\overline{X}_0$ y $\partial S_k(\overline{X})$ se hallan a una distancia positiva δ uno de la otra (véase el lema 7 en el p. 18.2).

$$\delta = \rho(\overline{X}_0, \partial S_k(\overline{X})) > 0. \tag{44.50}$$

Por ello todo conjunto D de diámetro $d(D)$ inferior a δ, que se corta con el conjunto $X_0 \subset \overline{X}_0$, se dispondrá íntegramente en $S_k(\overline{X}_0)$ (fig. 182). Efectivamente, si $d(D) < \delta$ y existe $x \in D \cap X_0$, entonces (véase (44.50)) $D \subset U(x, \delta) \subset S_k(\overline{X}_0)$, donde, como hasta ahora, $U(x, \delta)$ es el entorno esférico del punto x de radio δ.

Sea ahora $\tau = \{X_i\}$ una partición del conjunto X de finura $\delta_\tau < \delta$. Entonces para todo elemento X_i de esta partición cuya clausura se corta con el conjunto X_0, es decir, para cada $X_i \in \tau_0(X_0)$ tendremos $X_i \subset S_k(\overline{X}_0)$. Por esta razón

$$\bigcup_{X_i \in \tau_0(X_0)} X_i \subset S_k(\overline{X}_0).$$

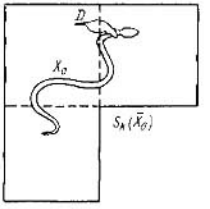


Fig. 182

Por consiguiente, en virtud de (44.49),

$$\sum_{X_i \in \tau_0(X_0)} \mu X_i = \mu \bigcup_{X_i \in \tau_0(X_0)} X_i \leqslant \mu S_k(\bar{X}_0) < \varepsilon. \square$$

Introduzcamos una designación más. Sea X un conjunto medible, sea $\tau = \{X_i\}_{i=1}^{n}$ cierta partición de este conjunto, $X_0 \subset X$. Para cualquier función f, definida en X, pongamos (véanse (44.46) y (44.47))

$$\sigma_{\tau(X_0)} = \sigma_{\tau(X_0)}(f; \xi^{(i)}, \ldots, \xi^{(i_0)}) = \sum_{X_i \in \tau(X_0)} f(\xi^{(i)}) \mu X_i. \qquad (44.51)$$

Esta inscripción significa que la sumación en el segundo miembro de la igualdad se realiza sólo según aquellos índices i, para los cuales $X_i \in \tau(X_0)$. Como siempre, $\xi^{(i)} \in X_i$. Para la simetría de la inscripción, las sumas integrales ordinarias de Riemann pueden, por analogía, anotarse en la forma

$$\sigma_\tau = \sum_{X_i \in \tau} f(\xi^{(i)}) \mu X_i.$$

En lugar del símbolo de sumación $\sum_{X_i \in \tau}$ a veces, para abreviar, se escribirá $\sum_{\tau}$. El límite de las sumas $\sigma_{\tau(X_0)}$ para $\delta_\tau \to 0$ se determina por analogía con el límite correspondiente de las sumas integrales σ_τ.

Teorema 5. *Supongamos que X es un conjunto del espacio R^n medible según Jordan y $\tau = \{X_i\}_{i=1}^{n}$ es la partición en este conjunto, $X_0 \subset X$ y $\mu X_0 = 0$. Si la función f está acotada en el conjunto X, entonces la integral de Riemann*

$$\int f(x)dX = \lim_{\delta_\tau \to 0} \sigma_\tau$$

existe cuando, y sólo cuando, existe el límite

$$\lim_{\delta_\tau \to 0} \sigma_{\tau(X_0)}.$$

Además, si el último límite existe, es igual a la integral $\int f(x)dX$.

Las sumas $\sigma_{\tau(X_0)}$ se llamarán, para abreviar, sumas integrales incompletas (correspondientes al conjunto X_0).

DEMOSTRACIÓN. Cualquiera que sea la partición $\tau = \{X_i\}_{i=1}^{n} = \{p$ del conjunto X, la clausura $\overline{X}_i$ de todo elemento X_i o bien no se corta con el conjunto X_0 y en este caso $X_i \in \tau(X_0)$ (véase (44.46)), o bien sí se corta (véase (44.47)) y en tal caso $X_i \in \tau_0(X_0)$. Por consiguiente, $\tau = \tau(X_0) \cup \tau_0(X_0)$, con la particularidad de que $\tau(X_0)$ y $\tau_0(X_0)$ no tienen elementos comunes.

Pongamos

$$\sigma_{\tau_0(X_0)} = \sum_{X_i \in \tau_0(X_0)} f(\xi^{(i)})\mu X_i, \quad \xi^{(i)} \in X_i.$$

Aquí la sumación en el segundo miembro se realiza según aquellos índices i, para los cuales $X_i \in \tau_0(X_0)$. Es evidente que para cualquier suma integral de Riemann se verifica la igualdad (véanse (44.42) y (44.51))

$$\sigma_\tau = \sigma_{\tau(X_0)} + \sigma_{\tau_0(X_0)}. \tag{44.52}$$

Por ser la función f acotada en X, existe tal constante $M > 0$, que para todo $x \in X$ se verifica la desigualdad $|f(x)| \leqslant M$. Por eso

$$|\sigma_{\tau_0(X_0)}| \leqslant \sum_{X_i \in \tau_0(X_0)} |f(\xi^{(i)})|\,\mu X_i \leqslant M \sum_{X_i \in \tau_0(X_0)} \mu X_i.$$

Dado que, de acuerdo con el lema 6, $\lim_{\delta_\tau \to 0} \sum_{X_i \in \tau_0(X_0)} \mu X_i = 0$, se tiene

$$\lim_{\delta_\tau \to 0} \sigma_{\tau_0(X_0)} = 0.$$

Debido a esto, de la igualdad (44.52) se desprende que la suma integral σ_τ y la suma integral incompleta $\sigma_{\tau(X_0)}$ simultáneamente o bien tienen los límites o no los tienen cuando $\delta_\tau \to 0$, con la particularidad de que si estos límites existen, son iguales. □

Del teorema demostrado se deduce que si una función está definida y es acotada en cierto conjunto medible X, entonces, al definir la integral como un límite de sumas integrales, podemos desechar en éstas todos los sumandos que corresponden a los elementos de la partición cuyas clausuras contienen puntos de frontera, pues el conjunto $X_0 = \partial X$ tiene medida cero (véase el teorema 1 en el p. 44.1).

Del teorema 5 proviene, además, que si una función f está definida y es continua en el conjunto medible X, la variación de sus valores en cierto conjunto $X_0 \subset X$ de medida cero que tiene por resultado la aparición de otra función, también acotada en X, no influye en la integrabilidad de la función, ni tampoco en el valor de la integral de la función, si existe. Esto se sigue directamente de lo que con la variación indicada de la función la suma $\sigma_{\tau(X_0)}$ queda inalterable, mientras que, en virtud del teorema 5, si su límite para $\delta_\tau \to 0$ existe, es igual a la integral $\int f(x)dX$:

$$\lim_{\delta_\tau \to 0} \sigma_{\tau(X_0)} = \int f(x)dX.$$

De esta observación se infiere, en particular, que la función f es integrable en el conjunto medible X cuando, y sólo cuando, en dicho conjunto sea integrable cualquier función que se obtiene de f por cambio arbitrario de sus valores en los puntos de frontera (es decir, en el conjunto $X \cap \delta X$) de un modo tal que los valores citados queden, sin embargo, acotados. Al realizarse esta operación, tampoco cambia el valor de la integral $\int f(x)dX$. Todo esto se debe a que la frontera de un conjunto medible y, por tanto, cualquier parte de ésta tienen medida cero.

De este modo, la integrabilidad y el valor de una integral de la función extendida al conjunto X no dependen de los valores de la función en los puntos de frontera del conjunto medible X, siempre que estos valores son acotados.

44.4. EXISTENCIA DE LA INTEGRAL

Como ejemplo más sencillo de función integrable según Riemann interviene la función numérica arbitraria f, definida en cierto conjunto $X \subset R^n$, cuya medida de Jordan es igual a cero: $\mu X = 0$. En este caso para cualquier partición $\tau = \{X_i\}_{i=1}^{i_0}$ del conjunto X tendremos $\mu X_i = 0$ para todo $i = 1, 2, \ldots, i_0$, eso, sea cual sea la elección de los puntos $\xi^{(i)} \in X_i$, obtendremos $f(\xi^{(i)})\mu X_i = 0$ y, por consiguiente, (véase (44.42)),

$$\sigma_\tau = \sigma_\tau(f; \xi^{(1)}, \ldots, \xi^{(i_0)}) = \sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i = 0.$$

De aquí, de acuerdo con la definición de la integral, ella, en este caso, existe y es igual a cero:

$$\int f(x)dX = \lim_{\delta_\tau \to 0} \sigma_\tau = 0.$$

Como la función f es arbitraria, puede ser, en particular, no acotada. En otras palabras, para su integrabilidad según Riemann en un conjunto arbitrario medible según Jordan, la condición de que dicha función sea acotada no es necesaria. Recordemos que para que una función sea integrable según Riemann en un segmento, la condición de acotar esta función era necesaria (véase el teorema 1 en el p. 27.2). Sin embargo, introducidas ciertas modificaciones, el teorema sobre la acotación de una función integrable resulta válido también para la integral que aquí se considera.

Demostremos previamente un lema.

Lema 7. *Supongamos que la función f está definida en un conjunto medible según Jordan X, $\tau = \{X_i\}_{i=1}^{i_0}$ es la partición de este conjunto y X^*, la reunión de todos los elementos de dicha partición que tienen medida positiva: $X^* = \bigcup_{\mu X_i > 0} X_i$.*

Si la función f no está acotada en el conjunto X^*, entonces, cualquiera que sea el número $M > 0$, se pueden escoger los puntos $\xi^{(i)} \in X_i$ de modo tal que sea válida la desigualdad

$$\left| \sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i \right| > M.$$

Corolario. *Sea f una función definida en el conjunto medible según Jordan X. Si para X existen las particiones tan pequeñas como se quiera, para las cuales la función f no está acotada en la unión de todos elementos de medida positiva, entonces f no será integrable en X.*

DEMOSTRACIÓN DEL LEMA. Por hipótesis del lema, el conjunto X^* es una reunión de los elementos X_i de medida positiva de la partición τ. Puesto que cualquier partición consta de un número finito de elementos, X^* es una suma finita de los conjuntos citados $X_i \in \tau$. Por ello, si la función f no está acotada en el conjunto X^*, tampoco está en cierto conjunto X_i de medida positiva. Sea este conjunto, para concretar, X_1. Debido a que la función f no está acotada en X_1, se puede escoger tal sucesión $\xi_n^{(1)} \in X_1$, $n = 1, 2, \ldots$, que se verificará la igualdad $\lim_{n \to \infty} f(\xi_n^{(1)}) = \infty$. Fijemos de tal o cual manera los puntos restantes $\xi^{(i)} \in X_i$ para $i = 2, 3, \ldots, i_0$.

Ya que la suma $\sum_{i=2}^{n} f(\xi^{(i)})\mu X_i$ es un número fijo y $\mu X_1 > 0$, en la suma

$$f(\xi_n^{(1)})\mu X_1 + \sum_{i=2}^{i_0} f(\xi^{(i)})\mu X_i$$

el primer sumando tiende al infinito cuando $n \to \infty$, mientras que el segundo sumando es una constante; por consiguiente

$$\lim_{n \to \infty} \left| f(\xi^{(1)})\mu X_1 + \sum_{i=2}^{i_0} f(\xi^{(i)})\mu X_i \right| = +\infty.$$

Por eso, para cualquier número $M > 0$ podemos elegir tal número $n_0 = n_0(M)$ que se verifique la desigualdad

$$\left| f(\xi_{n_0}^{(1)})\,\mu X_1 + \sum_{i=2}^{i_0} f(\xi^{(i)})\mu X_i \right| > M. \quad \square$$

DEMOSTRACIÓN DEL LEMA. Si la función f es integrable en el conjunto X, es decir, si existe el límite

$$\lim_{\delta_\tau \to 0} \sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i = \int f(x)dX,$$

entonces para todo $\varepsilon > 0$, por ejemplo, para $\varepsilon = 1$, existe tal $\delta_0 > 0$ que cualesquiera que sean las particiones $\tau = \{X_i\}_{i=1}^{i_0}$ del conjunto X de finura $\delta < \delta_0$, no importa cual sea la elección de los puntos $\xi^{(i)} \in X_i \in \tau$ se verifica la desigualdad

$$\left| \sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i - \int f(x)dX \right| < 1$$

y, por consiguiente, la desigualdad

$$\int f(x)dX - 1 < \sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i < \int f(x)dX + 1. \tag{44.53}$$

En cambio, si la función f satisface las condiciones del corolario, entonces para el conjunto X existe la partición τ de finura $\delta_\tau < \delta_0$, para la cual la función f no está acotada en la unión de todos los elementos de medida positiva de esta partición. En este caso, conforme al lema 7, la suma $\sum_{i=1}^{i_0} f(\xi^{(i)})\mu X_i$ puede hacerse, debido a la elección adecuada de los puntos $\xi^{(i)} \in X_i \in \tau$, tan grande como se quiera en su magnitud absoluta. Por ello, tal función no puede ser integrable, pues para ella no se cumple la condición (44.53). □

Mostremos ahora que si menospreciamos el conjunto de medida cero, toda función integrable será acotada.

Teorema 6. *Si una función f es integrable en el conjunto X, existe tal conjunto de medida cero $X_0 \subset X$: $\mu X_0 = 0$, que la función f esté acotada en $X \setminus X_0$.*

DEMOSTRACIÓN. Supongamos que la función f es integrable en X y el conjunto X_0, citado en el teorema, no existe. Tomemos cualquier $\delta_0 > 0$ y una partición τ del conjunto X de finura $\delta_\tau < \delta_0$. Designemos con X^* la reunión de todos los elementos de medida positiva. El conjunto $X \setminus X^*$ es, pues, la reunión de un número finito de conjuntos $X_i \in \tau$ de medida cero, por lo cual el mismo conjunto tiene medida cero: $\mu(X \setminus X^*) = 0$. Consecuentemente, según la hipótesis aceptada, la función f no está acotada en el conjunto X^*. De aquí, de acuerdo con el corolario del lema 7, llegamos a que la función f no es integrable. □

Probemos ahora que para una clase importante de los conjuntos abiertos medibles según Jordan el teorema sobre la acotación de las funciones integrables queda enteramente en vigor. Para demostrarlo nos hará falta un lema geométrico.

Lema 8. *La intersección no vacía de un cubo cerrado n-dimensional con un conjunto abierto de un espacio n-dimensional tiene la medida inferior de Jordan positiva.*

Corolario. *Para cualquier conjunto abierto medible según Jordan existen particiones tan pequeñas como se quiera, cuyos elementos tienen todos medida positiva.*

DEMOSTRACIÓN DEL LEMA. Supongamos que Q es un cubo n-dimensional y sea G un conjunto abierto del espacio R^n y $Q \cap G \neq \varnothing$. Cualquiera que sea el punto $x \in Q \cap G$, por ser abierto el conjunto G, existe tal entorno de dicho punto $U(x)$ (por ejemplo, el propio conjunto G) que

$$U(x) \subset G. \tag{44.54}$$

No es difícil convencerse de que en el conjunto $U(x)$ siempre hay un punto interior y del cubo Q. Efectivamente, puede ocurrir que el mismo punto x sea interior para el cubo Q, y en tal caso podemos tomar $y = x$. En cambio, si x es un punto de frontera del cubo Q, será punto de frontera también para el conjunto de los puntos interiores del cubo. Por esta razón su entorno $U(x)$ contiene, a ciencia cierta, el punto interior y del cubo Q (fig. 183). En virtud de la definición de punto interior (véase el p. 18.2), existe tal entorno suyo $V(y)$ que

$$V(y) \subset Q. \tag{44.55}$$

En vista de (44.54) y (44.55), son lícitas las inclusiones

$$U(x) \cap V(y) \subset U(x) \subset G, \quad U(x) \cap V(y) \subset V(y) \subset Q;$$

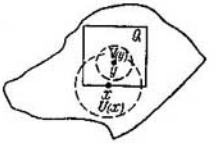


Fig. 183

y por ende

$$U(x) \cap V(y) \subset Q \cap G. \tag{44.56}$$

Dado que $y \in U(x)$ e $y \in V(y)$, la intersección $U(x) \cap V(y)$ no es vacía, pues contiene en todo caso el punto y. Luego, siendo la misma una intersección de dos conjuntos abiertos, es también abierta, razón por la cual (véase el p. 44.1)

$$\mu_*[U(x) \cap V(y)] > 0.$$

En virtud de la propiedad de monotonía de la medida inferior (véase el lema 1 en el p. 44.1), de (44.56) tenemos

$$\mu_*[U(x) \cap V(y)] \leqslant \mu_*(Q \cap G).$$

De las últimas dos desigualdades se pone de manifiesto que $\mu_*(Q \cap G) > 0$. □

DEMOSTRACIÓN DEL COROLARIO. Sea G un conjunto medible abierto en R^n. Fijemos una partición del espacio R^n en cubos de cierto rango k. El conjunto de cubos Q de este rango que tienen una intersección no vacía con el conjunto G, es finito, pues el conjunto G está acotado, debido a su mensurabilidad. Numeremos todos los cubos mencionados: $Q_1, Q_2, \ldots, Q_{i_0}$. Los conjuntos $X_i = Q_i \cap G = \varnothing$, $i = 1, 2, \ldots, i_0$, son medibles y forman la partición $\tau = \{X_i\}_{i=1}^{i_0}$ del conjunto G. En efecto, por una parte $X_i = Q_i \cap G \subset G$, por consiguiente, $\bigcup_{i=1}^{i_0} X_i \subset G$, pero, por otra parte, todo punto $x \in G$, al igual que cualquier otro punto del espacio R^n, pertenece por lo menos a un cubo Q de rango k : $x \in Q \cap G$. En este caso $Q \cap G \in \tau$, es decir, para cierto i se tiene $Q = Q_i$, por lo cual $x \in Q_i \cap X = X_i \subset \bigcup_{i=1}^{i_0} X_i$. De este modo, $\bigcup_{i=1}^{i_0} X_i = G$.

Luego, $X_i \cap X_j \subset Q_i \cap Q_j$. Si la intersección $Q_i \cap Q_j$ no es vacía, representa un cubo de dimensión inferior a n, y, por consiguiente, es la gráfica de una función continua (incluso lineal) en un compacto. Por esta razón su medida es igual a cero: $\mu(Q_i \cap Q_j) = 0$, de donde también $\mu(X_i \cap X_j) = 0$, $i \neq j$. Por fin, de conformidad con el lema 8, $\mu X_i > 0$.

Existen, evidentemente, las particiones del tipo indicado tan pequeñas como se quiera. En efecto, cualquiera que sea $\delta > 0$, basta elegir tal rango k que sea $10^{-k}\sqrt{n} <$

$< \delta$ ($d(Q) = 10^{-k}\sqrt{n}$ es el diámetro del cubo de rango k) y entonces

$$d(X_i) = d(Q_i \cap G) \leqslant d(Q_i) = 10^{-k}\sqrt{n} < \delta, \quad i = 1, 2, \ldots, i_0,$$

por lo cual $\delta_\tau < \delta$. □

Teorema 7. *Si una función es integrable en un conjunto abierto, está acotada.*

DEMOSTRACIÓN. Supongamos que la función f es integrable en un conjunto medible G. Según la definición de integral, el conjunto G es medible según Jordan y debido al corolario del lema 8, existen particiones, tan pequeñas como se quiera, de dicho conjunto, todos los elementos de las cuales tienen medida positiva. Es evidente que, en virtud del lema 8 para las particiones construidas al demostrar el corolario del lema citado, la reunión de todos sus elementos de medida positiva coincide con el mismo conjunto G. Si la función f fuera no acotada en G, entonces, de acuerdo con el corolario del lema 7, no sería integrable.

OBSERVACIÓN. Según se ve de la demostración aducida del teorema 7, el carácter abierto del conjunto G se ha requerido sólo para mostrar que existen sus particiones, tan pequeñas como se quiera, cuyos elementos tienen todos medida positiva. De este modo, para todos los conjuntos que poseen esta propiedad la integrabilidad de las funciones en ellos conlleva a su carácter acotado.

Podemos convencernos con facilidad de que la clausura G de todo conjunto abierto medible G también tiene particiones, tan pequeñas como se quiera, todos los elementos de las cuales *tienen* medida positiva. Efectivamente, basta tomar de nuevo todos los cubos Q_i de rango k que hacen con G una intersección no vacía. En este caso con mayor razón tendrán no una intersección vacía con la clausura $\bar{G}$ del conjunto G: $Q_i \cap \bar{G} \supset Q_i \cap G = \varnothing$. Además, como $S_k(G)$ es un conjunto cerrado y $G \subset S_k(G)$, entonces $\bar{G} \subset S_k(G)$. Por lo tanto, si ponemos $X_i = Q_i \cap \bar{G}$, donde $Q_i \cap G \neq \varnothing$, la partición $\tau = \{X_i\}$ forma el recubrimiento de la clausura $\bar{G}$ del conjunto G, pues el poliedro $S_k(G)$ se compone sólo de los cubos indicados Q_i y, según el lema 8,

$$\mu X_i = \mu(Q_i \cap \bar{G}) \geqslant \mu(Q_i \cap G) > 0.$$

Ejercicio 6. Constrúyase un ejemplo de función no acotada e integrable en un conjunto *de medida positiva*

Si la función f está acotada en un conjunto medible, se pueden determinar, igual que en el caso unidimensional, las sumas superiores e inferiores de Darboux.

Definición 5. *Sea f una función acotada en el conjunto medible según Jordan X, sea $\tau = \{X_i\}_{i=1}^{i_0}$ la partición del conjunto X,*

$$m_i = \inf_{x \in X_i} f(x), \quad M_i = \sup_{x \in X_i} f(x), \quad i = 1, 2, \ldots, i_0.$$

En tal caso las sumas

$$s_\tau = \sum_{i=1}^{i_0} m_i \mu X_i, \quad S_\tau = \sum_{i=1}^{i_0} M_i \mu X_i$$

llevan los nombres respectivos de sumas inferiores y superiores de Darboux.

Para las sumas de Darboux y sumas integrales de Riemann son válidas las siguientes desigualdades evidentes

$$s_\tau \leqslant \sigma_\tau \leqslant S_\tau.$$

Lo mismo que para la función de una sola variable, para cualesquiera dos particiones τ_1 y τ_2 es lícita la desigualdad

$$s_{\tau_1} \leqslant S_{\tau_2}.$$

Teorema 8. *Para que la función f, acotada en un conjunto medible según Jordan $X \subset R^n$, sea integrable según Riemann en dicho conjunto, es necesario y suficiente que*

$$\lim_{\delta_\tau \to 0} (S_\tau - s_\tau) = 0. \tag{44.57}$$

Si se cumplen estas condiciones, se tiene

$$\lim_{\delta_\tau \to 0} S_\tau = \lim_{\delta_\tau \to 0} s_\tau = \int f(x)dX. \tag{44.58}$$

La condición (44.57) es equivalente a la siguiente

$$\lim_{\delta_\tau \to 0} \sum_{i=1}^{i_0} \omega(f; X_i)\mu X_i = 0, \tag{44.59}$$

donde $\omega(f; X_i)$ es la oscilación de la función f en el conjunto $X_i \in \tau = \{X_i\}$.

La demostración de este teorema es análoga a la que se realiza en el caso unidimensional y se recomienda que esta demostración quede a cargo del propio lector.

Ejercicio 7. Enúnciense las definiciones de los límites (44.57) — (44.59) con ayuda de las sucesiones y empleando el "lenguaje $\varepsilon - \delta$".

Teorema 9. *Si una función es continua en un compacto medible según Jordan, será integrable en éste.*

DEMOSTRACIÓN. Sea X un compacto medible, $X \subset R^n$, y sea f una función continua en dicho compacto. Toda función, continua en un compacto, está acotada (véase el p. 19.6) y es uniformemente continua (véase el p. 19.7) en él. Por ello, aquí también la demostración es la misma que en el caso unidimensional (véase el p. 27.5): se obtiene con facilidad la estimación

$$\sum_{i=1}^{i_0} \omega(f; X_i)\mu X_i \leqslant \omega(\delta_\tau; f)\mu X,$$

donde $\omega(\delta, f)$ es el módulo de continuidad de la función f. De esta estimación se deduce directamente el cumplimiento de la condición (44.59), por lo cual, conforme al teorema 8, tiene lugar también la integrabilidad de la función f. □

44.5°. SOBRE LA INTEGRABILIDAD DE LAS FUNCIONES DISCONTINUAS

La continuidad de una función no es una condición necesaria de integrabilidad: existen también funciones integrables discontinuas. Una clase bastante amplia de las funciones integrables discontinuas se determina por el siguiente teorema.

Teorema 10. *Si una función está acotada en un compacto medible según Jordan y el conjunto de puntos de discontinuidad de la función citada tiene medida de Jordan cero la función es integrable según Riemann.*

DEMOSTRACIÓN. Supongamos que la función está definida y acotada en un compacto, esto es, en un conjunto cerrado acotado $X \subset R^n$, con la particularidad de que el compacto X es medible según Jordan. Siendo f acotada en X, existe tal constante $M > 0$, que para todo $x \in X$ se cumple la desigualdad

$$|f(x)| \leqslant M. \tag{44.60}$$

Sea X_0 un conjunto de los puntos de discontinuidad de la función f. De conformidad con la hipótesis del teorema, $\mu X_0 = 0$, razón por la cual para todo $\varepsilon > 0$ fijo existe tal rango k que

$$\mu S_k(X_0) < \frac{\varepsilon}{3^n 4M}. \tag{44.61}$$

Esto es una consecuencia de que en el caso dado, con arreglo a la definición de medida, $\lim_{k \to +\infty} \mu S_k(X_0) = 0$. Supongamos que el poliedro $S_k(X_0)$ se compone de los cubos $Q_1, Q_2, \ldots, Q_l$. Denotemos con P_j un cubo que se obtiene de Q_j con ayuda de la transformación de semejanza con centro en el centro del cubo Q_j y la razón de semejanza igual a tres; en este caso

$$\mu P_j = 3^n \mu Q_j, \quad j = 1, 2, \ldots, l. \tag{44.62}$$

Pongamos $P = \bigcup_{j=1}^{s} P_j$. En virtud de las desigualdades (44.61) y (44.62), tenemos

$$\mu P = \mu \bigcup_{j=1}^{s} P_j \leqslant \sum_{j=1}^{s} \mu P_j = \sum_{j=1}^{s} 3^n \mu Q_j = 3^n \mu S_k(X_0) < \frac{\varepsilon}{4M}. \tag{44.63}$$

Cabe notar que el conjunto P se obtiene de $S_k(X_0)$, orlando el último con una franja de cubos cuyas aristas son de longitud 10^{-k}, por lo cual todo conjunto A de diámetro $d(A)$ inferior a 10^{-k}, que se corta con el conjunto $S_k(X_0)$, está contenido en P (fig. 184):

$$d(A) < 10^{-k}, \quad A \cap S_k(X_0) = \varnothing \Rightarrow A \subset P. \tag{44.64}$$

Designemos ahora con G el conjunto de puntos interiores del poliedro $S_k(X_0)$. Evidentemente, G es un conjunto abierto y como, según la hipótesis, del teorema, X es cerrado, entonces el conjunto $F = X \setminus G$ es también cerrado y, además por ser X acotado, el conjunto F es también acotado, razón por la cual F es un compacto. Luego, el conjunto X_0 se dispone en el interior del poliedro $S_k(X_0)$, es decir, $X_0 \subset G$ (según se ha observado anteriormente, véase el p. 44.1, esto tiene lugar, en general, para cualquier conjunto X y se deduce de la definición del poliedro $S_k(X)$). De aquí se pone claro que la función f es continua en el compacto F y, como, además, el conjunto F es medible en su calidad de diferencia entre dos conjuntos medibles X y G, entonces, de acuerdo con el teorema 9, la función f es integrable en F.

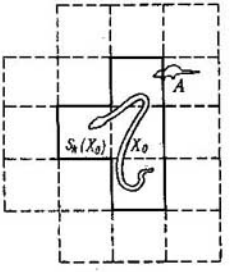


Fig. 184

Por eso, para $\varepsilon > 0$ elegido más arriba, existe tal $\delta > 0$, que para toda partición τ_F del conjunto F de finura $\delta_{\tau_F} < \delta$ se verifica la desigualdad

$$S_{\tau_F} - s_{\tau_F} < \frac{\varepsilon}{2}, \tag{44.65}$$

donde S_{τ_F} y s_{τ_F} son las sumas superiores e inferiores de Darboux de la función f, correspondientes a la partición τ_F del conjunto F.

Sea

$$\delta_0 = \min\{10^{-k}, \delta\}, \tag{44.66}$$

$\tau = \{X_i\}_{i=1}^{i_0}$ una partición del conjunto X de finura $\delta_\tau < \delta_0$. Es evidente que $\tau_F \overset{\text{def}}{=} \{X_i \cap F\}$, donde $X_i \cap F \neq \varnothing$, es una partición del conjunto F de finura $\delta_{\tau_F} \leqslant \delta_\tau < \delta_0$, y, por esta razón, en virtud de (44.66), para τ_F se verifica la desigualdad (44.65).

Pongamos

$$M_i = \sup_{x \in X_i} f(x), \quad m_i = \inf_{x \in X_i} f(x),$$

$$S_\tau = \sum_{i=1}^{i_0} M_i \mu X_i, \; s_\tau = \sum_{i=1}^{i_0} m_i \mu X_i,$$

$$M_i' = \sup_{x \in X_i \cap F} f(x), \; m_i' = \inf_{x \in X_i \cap F} f(x),$$

$$S_{\tau_F} = \sum_{X_i \cap F \neq \varnothing} M_i' \mu(X_i \cap F), \; s_{\tau_F} = \sum_{X_i \cap F \neq \varnothing} m_i' \mu(X_i \cap F).$$

Todo conjunto $X_i \in \tau$ se interseca con G o no se interseca. Si no se interseca, es decir, si $X_i \cap G = \varnothing$, se tiene $X_i \subset F$ y para tales índices i tenemos $M_i = M_i'$, $m_i = m_i'$, $X_i \cap F = X_i$.

Dado que $X_i \neq \varnothing$ y $X_i \subset X = F \cup G$, de $X_i \cap G = \varnothing$ se deduce que $X_i \subset F$ y, por consiguiente, $X_i \cap F \neq \varnothing$. Por ello, al observar que en las sumas que vienen

abajo todos los sumandos son no negativos, obtendremos

$$\sum_{X_i \cap G = \varnothing} (M_i - m_i)\mu X_i = \sum_{X_i \cap G = \varnothing} (M_i' - m_i')\mu X_i \leqslant \sum_{X_i \cap F \neq \varnothing} (M_i' - m_i')\mu(X_i \cap F) =$$

$$= S_{\tau_F} - s_{\tau_F} < \frac{\varepsilon}{2}. \quad (44.67)$$

Si $X_i \cap G \neq \varnothing$, en virtud de (44.64) y (44.66), $X_i \subset P$ y por esta razón para estos índices i (véase, además, (44.63)) tenemos

$$\sum_{X_i \cap G \neq \varnothing} \mu X_i = \mu \bigcup_{X_i \cap G \neq \varnothing} X_i \leqslant \mu P < \frac{\varepsilon}{4M}. \quad (44.68)$$

Al hacer uso de las desigualdades obvias $|m_i| \leqslant M$, $|M_i| \leqslant M$, $i = 1, 2, \ldots, i_0$, que se desprenden directamente de (44.60) y al aplicar la desigualdad (44.68), tendremos

$$\sum_{X_i \cap G \neq \varnothing} (M_i - m_i)\mu X_i \leqslant \sum_{X_i \cap G \neq \varnothing} [|M_i| + |m_i|]\mu X_i \leqslant 2M \sum_{X_i \cap G \neq \varnothing} \mu X_i < \frac{\varepsilon}{2}. \quad (44.69)$$

De (44.67) y (44.69) se deduce que

$$S_\tau - s_\tau = \sum_{i=1}^{i_0} (M_i - m_i)\mu X_i =$$

$$= \sum_{X_i \cap G \neq \varnothing} (M_i - m_i)\mu X_i + \sum_{X_i \cap G = \varnothing} (M_i - m_i)\mu X_i < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

De aquí, de acuerdo con el teorema 8, se deduce la integrabilidad de la función f en el conjunto X. □

44.6. PROPIEDADES DE LA INTEGRAL MÚLTIPLE

Consideraremos aquí las propiedades de la integral múltiple análogas a las de la función de una sola variable por un segmento. Recordemos que la integrabilidad de una función (según Riemann) en cierto conjunto presupone la mensurabilidad de éste según Jordan.

1°. *Sea X un conjunto medible; en este caso* $\int dX = \mu X$.

Efectivamente, el integrando es idénticamente igual a la unidad. Por ello, si $\tau = \{X_i\}_{i=1}^{i_0}$ es cierta partición del conjunto X, entonces (véase (44.41))

$$\int dX = \lim_{\delta_\tau \to 0} \sum_{i=1}^{i_0} \mu X_i = \mu X. \ \square$$

2°. *Sean X y X^* unos conjuntos medibles, $X^* \subset X$ y sea f una función acotada e integrable en X; será también integrable en X^*.*

En efecto, el conjunto $X^{**} = X \setminus X^*$ es también medible, siendo una diferencia entre dos conjuntos medibles. Sean $\tau^* = \{X_i^*\}$ una partición del conjunto X^* de finura δ_{τ^*} y $\tau^{**} = \{X_j^{**}\}$ una partición del conjunto X^{**} de finura $\delta_{\tau^{**}} < \delta_{\tau^*}$. Entonces, $\tau = \{X_i^*, X_j^{**}\}$ será la partición del conjunto X de finura $\delta_\tau = \delta_{\tau^*}$. Si

$$\omega_\tau = \sum_{\tau^*} \omega(f, X_i^*)\,\mu X_i^* + \sum_{\tau^{**}} \omega(f, X_j^{**})\,\mu X_j^{**}$$

y

$$\omega_{\tau^*} = \sum_{\tau^*} \omega(f, X_i^*)\,\mu X_i^*,$$

entonces, evidentemente, $0 \leqslant \omega_{\tau^*} \leqslant \omega_\tau$. Pero $\lim_{\delta_\tau \to 0} \omega_\tau = 0$, por lo cual $\lim_{\delta_{\tau^*} \to 0} \omega_{\tau^*} = 0$, de donde proviene la integrabilidad de la función f en el conjunto X^* (véase (44.59)). □

3°. **Aditividad de la integral por los conjuntos.** *Si X y X'' son unos conjuntos medibles, $X = X' \cup X''$, $X' \cap X'' = \varnothing$, y la función f está acotada y es integrable en el conjunto X, entonces las integrales $\int f(x)dX'$ y $\int f(x)dX''$ existen y*

$$\int f(x)dX = \int f(x)dX' + \int f(x)dX''. \tag{44.70}$$

Ya que la existencia de las integrales $\int f(x)dX'$ y $\int f(x)dX''$ proviene de la propiedad 2°, nos queda demostrar sólo la fórmula (44.70). Sean $\tau' = \{X_i'\}$ y $\tau'' = \{X_j''\}$ las particiones de los conjuntos respectivos X' y X''. En este caso $\tau = \{X_i', X_j''\}$ es la partición del conjunto X y su finura es igual a las máxima de las finuras de las particiones $\delta_{\tau'}$ y $\delta_{\tau''}$: $\delta_\tau = \text{máx}\,[\delta_{\tau'}, \delta_{\tau''}]$.

Sea $\xi^{(i)} \in X_i'$, $\eta^{(j)} \in X_j''$,

$$\sigma_{\tau'} = \sum_{\tau'} f(\xi^{(i)})\,\mu X_i', \quad \sigma_{\tau''} = \sum_{\tau''} f(\eta^{(j)})\,\mu X_j'',$$

$$\sigma_\tau = \sigma_{\tau'} + \sigma_{\tau''}. \tag{44.71}$$

Por ser f integrable en los conjuntos X, X' y X'', tenemos

$$\lim_{\delta_\tau \to 0} \sigma_\tau = \int f(x)dX, \quad \lim_{\delta_{\tau'} \to 0} \sigma_{\tau'} = \int f(x)dX', \quad \lim_{\delta_{\tau''} \to 0} \sigma_{\tau''} = \int f(x)dX'',$$

Por ello, pasando al límite en la igualdad (44.71) para $\delta_\tau \to 0$, obtendremos (44.70). □

OBSERVACIÓN. Conviene fijar la atención en la siguiente circunstancia: puede suceder que la función f esté definida en el conjunto $X = X' \cup X''$, donde X' y X'' son conjuntos medibles, $X' \cap X'' = \varnothing$, las integrales $\int f(x)dX'$ y $\int f(x)dX''$ existen, mientras que la integral $\int f(x)dX$ no existe.

Aclaremos lo dicho con un ejemplo. Sean (r, φ) las coordenadas polares de un punto en un plano

$$f(r, \varphi) = \begin{cases} 0, & \text{si } r < 1, \\ 1/\varphi, & \text{si } r = 1,\ 0 < \varphi \leqslant 2\pi, \end{cases}$$

sea $X' = \{(r, \varphi): r < 1\}$ un círculo abierto y $X'' = \{(r, \varphi): r = 1\}$, una circunferencia. Evidentemente, $\mu X'' = 0$, por lo cual, pese a que la función f no está acotada en X'', es integrable y $\int f(r, \varphi) dX'' = 0$. Existe también la integral $\int f(r, \varphi) dX' = 0$. No obstante, la integral $\int f(r, \varphi)\, dX$ referida a un círculo cerrado $X = X' \cup X''$ no existe. Efectivamente, el conjunto X representa la clausura de una región, razón por la cual dicho conjunto tiene particiones, tan pequeñas como se quiera, todos los elementos de las cuales cuentan con una medida positiva. Por lo tanto (véase la observación al teorema 7), toda función integrable en X está acotada, en tanto que la función dada f no está acotada y no es, por ende, integrable.

Es importante observar que la situación semejante no puede tener lugar para las funciones acotadas: si la función f está acotada y es integrable en los conjuntos medibles X' y X'', $X' \cap X'' = \varnothing$, entonces será integrable también en el conjunto $X = X' \cup X''$, con la particularidad de que queda lícita la fórmula (44.70). Esto será demostrado en el p. 44.7*.

Sólo diremos que en el caso cuando uno de los conjuntos, X' o X'', tiene medida cero, la integrabilidad de la función acotada f en su reunión, si se presupone la integrabilidad de dicha función en cada uno de los conjuntos, se obtiene por repitición textual de los razonamientos aducidos en la demostración del teorema 10. En efecto, sea f integrable y acotada en los conjuntos medibles X' y X'', $\mu X' = 0$, $X = X' \cup X''$. En este caso si construimos, al igual que en la demostración aducida, un conjunto $G \supset X'$ (el conjunto X' desempeña aquí el papel del conjunto X_0 del teorema 10) y ponemos $F = X \setminus G$, entonces tendremos $F \subset X''$ y, por lo tanto, en virtud de la propiedad 2° de las integrales, la función f resultará integrable en el conjunto F, de donde se desprende, igual que antes, su integrabilidad en el conjunto X, y, por ende, en virtud de la propiedad 3°, la validez de la fórmula (44.70), donde $\int f(x) dX' = 0$.

Del modo semejante se demuestra la afirmación general, pero el procedimiento para ello resulta ser más complicado.

4°. **Linealidad de la integral.** *Si las funciones f_1 y f_2 son integrables en el conjunto X, para cualesquiera números λ_1 y λ_2 existe la integral $\int [\lambda_1 f_1(x) + \lambda_2 f_2(x)]\, dX$ y se verifica la igualdad*

$$\int [\lambda_1 f_1(x) + \lambda_2 f_2(x)]\, dX = \lambda_1 \int f_1(x)\, dX + \lambda_2 \int f_2(x)\, dX.$$

5°. *Si las funciones f y g son integrables y están acotadas en cierto conjunto, entonces su producto y la razón f/g (para $\inf_X |g| > 0$) son integrables en dicho conjunto.*

6°. **Integración de las desigualdades.** *Si las funciones f y g son integrables en el conjunto X y si para todo $x \in X$ se verifica la desigualdad $f(x) \leqslant g(x)$, entonces* $\int f(x) dX \leqslant \int g(x) dX$.

7°. *Si la función f es integrable y está acotada en un conjunto X, el valor absoluto de la función $|f|$ es también integrable en el conjunto, con la particularidad de que* $\left|\int f(x)dX\right| \leqslant \int |f(x)|\,dX$.

La demostración de las propiedades 4°, 5°, 6°, 7° se efectúa por analogía completa con el caso unidimensional (véase el p. 28.1).

8°. **Monotonía de la integral de las funciones no negativas extendida a los conjuntos.** *Si X y X^* son unos conjuntos medibles, $X^* \subset X$, y la función f es no negativa, acotada e integrable en X, entonces*

$$\int f(x)dX^* \leqslant \int f(x)dX. \tag{44.72}$$

En efecto, en vista de las propiedades 2° y 3°, las integrales $\int f(x)dX^*$ y $\int f(x)d(X \setminus X^*)$ existen y

$$\int f(x)dX = \int f(x)dX^* + \int f(x)d(E \setminus X^*).$$

Como $f(x) \geqslant 0$, se tiene, en virtud de la propiedad 6°, $\int f(x)d(X \setminus X^*) \geqslant 0$, de lo que proviene la desigualdad (44.72). □

9°. *Supongamos que la función f es integrable y no negativa en un conjunto abierto medible G, $x^{(0)} \in G$ y la función f es continua en el punto $x^{(0)}$, siendo $f(x^{(0)}) > 0$. En este caso*

$$\int f(x)dG > 0. \tag{44.73}$$

En efecto, por ser la función f continua en el punto $x^{(0)}$, para todo $\varepsilon > 0$ existe tal entorno $U = U(x^{(0)})$ de este punto que con cualquiera $x \in U$ se verifica la desigualdad $f(x^{(0)}) - \varepsilon < f(x) < f(x^{(0)}) + \varepsilon$. Además, debido a que el conjunto G es abierto, el entorno U siempre puede elegirse de modo tal que sea $U \subset G$.

Al escoger $\varepsilon = \dfrac{f(x^{(0)})}{2}$, obtendremos para éste tal entorno U que para cualesquiera x, pertenecientes a este entorno, tendremos $f(x) > \dfrac{f(x^{(0)})}{2}$. De aquí, aplicando sucesivamente las propiedades 8°, 6° y 1°, llegamos a que

$$\int f(x)dG \geqslant \int f(x)dU \geqslant \frac{f(x^{(0)})}{2} \int dU = \frac{f(x^{(0)})}{2}\,\mu U > 0,$$

pues $\mu U > 0$, como medida de cualquier conjunto abierto. □

Demos a conocer un corolario directo de la propiedad 9°.

Corolario. *Si la función f es continua, integrable y no negativa en un conjunto abierto medible G y no es idénticamente igual a cero, entonces* $\int f(x)dG > 0$.

10°. **Aditividad total de la integral extendida a los conjuntos.** *Supongamos que la función f está acotada y es integrable en el conjunto X, mientras que $\{X_k\}$, $k = 1, 2, \ldots$, es la sucesión de tales conjuntos medibles $X_k \subset X$, que*

$$\lim_{k \to +\infty} \mu X_k = \mu X.^{*)} \tag{44.74}$$

En este caso

$$\lim_{k \to +\infty} \int f(x)dX_k = \int f(x)dX. \tag{44.75}$$

Por ser la integral aditiva, tenemos

$$\int f(x)dX - \int f(x)dX_k = \int f(x)d(X \setminus X_k).$$

Como, según la hipótesis, f es una función acotada, es decir, existe tal constante $M > 0$ que $|f(x)| \leqslant M$ para todo $x \in X$, entonces

$$\left|\int f(x)dX - \int f(x)dX_k\right| = \left|\int f(x)d(X \setminus X_k)\right| \leqslant$$

$$\leqslant \int |f(x)|\, d(X \setminus X_k) \leqslant M \int d(X \setminus X_k) = M\mu(X \setminus X_k).$$

Según la propiedad de aditividad de la medida tenemos $\mu(X \setminus X_k) = \mu X - \mu X_k$, por consiguiente

$$\left|\int f(x)dX - \int f(x)dX_k\right| \leqslant M(\mu X - \mu X_k).$$

De aquí, en virtud de (44.74) proviene precisamente (44.75). □

11°. **Teorema del valor medio.** *Supongamos que las funciones f y g están acotadas y son integrables en el conjunto X. Si la función g no cambia de signo en X y $m \leqslant f(x) \leqslant M$, $x \in X$, entonces existe tal número λ, $m \leqslant \lambda \leqslant M$, que*

$$\int f(x)g(x)dX = \lambda \int g(x)dX.$$

Corolario. *Sea X un conjunto medible linealmente conexo o una clausura de un conjunto linealmente conexo. En este caso, si la función f está acotada, es integrable y continua en X, existe tal punto $\xi \in X$, que $\int f(x)dX = f(\xi)\mu X$.*

El teorema sobre el valor medio se demuestra por analogía completa con el caso unidimensional (véase el p. 28.2). Para obtener el corolario, se debe usar el teorema sobre los valores intermedios de la función continua en un conjunto linealmente conexo o en la clausura de éste (véase el p. 19.6).

44.7*. CRITERIOS DE INTEGRABILIDAD DE LAS FUNCIONES DE RIEMANN Y DARBOUX Y LOS COROLARIOS

Supongamos que la función f está definida y acotada en un conjunto medible según Jordan X, $\tau = \{X_i\}_{i=1}^{n} = \{p$ es la partición del conjunto, $m_i = \inf_X f$,

*) Con las sucesiones de conjuntos medibles, que poseen la propiedad (44.74), ya nos encontramos, véase, por ejemplo, el teorema 2 en el p. 31.2.

$M_i = \sup_X f, s_\tau = \sum_{i=1}^{i_0} m_i \mu X_i, S_\tau = \sum_{i=1}^{i_0} M_i \mu X_i$ son las sumas de Darboux, inferior y superior, correspondiente a la partición τ. Pongamos

$$I_* = \sup_\tau s_\tau, I^* = \inf_\tau S_\tau; \tag{44.76}$$

I_* e I^* se llaman *integrales de Darboux inferior y superior*, respectivamente, de la función f. Resulta que las integrales inferior y superior de Darboux representan no sólo las cotas superior e inferior de las sumas integrales de Darboux, sino también su límite a condición de que la finura de las particiones tiende a cero.

Teorema 11. *Si la función f está acotada en un conjunto medible según Jordan X, resulta*

$$I_* = \lim_{\delta_\tau \to 0} s_\tau, \quad I^* = \lim_{\delta_\tau \to 0} S_\tau.$$

DEMOSTRACIÓN. Demostremos la validez de la primera fórmula (la segunda se demuestra de modo análogo). Supongamos que $|f(x)| \leqslant M$, $x \in X$, y está dado $\varepsilon > 0$. En vista de la definición (44.76), existe tal partición $\tau^* = \{X_i^*\}$ del conjunto X que

$$s_{\tau^*} > I_* - \frac{\varepsilon}{3}. \tag{44.77}$$

Aquí $s_{\tau^*} = \sum_{i=1}^{i_0} m_i^* \mu X_i^*$, $m_i^* = \inf_{X_i^*} f$, $i = 1, 2, \ldots, i_0$. Sea

$$X_0 = \bigcup_{i=1}^{i_0} \partial X_i^*. \tag{44.78}$$

Por cuanto cada conjunto X_i^* es medible, $\mu \partial X_i^* = 0$; por ello $\mu X_0 = 0$. Por consiguiente, existe tal rango $k = k(\varepsilon)$, que

$$\mu S_k(X_0) < \frac{\varepsilon}{3^{n+1} M}. \tag{44.79}$$

Probemos que para cualquier partición $\tau = \{X_j\}_{j=1}^{j_0}$ del conjunto X de finura $\delta_\tau < 10^{-k}$ se verifica la igualdad

$$I_* - \varepsilon < s_\tau \leqslant I_*. \tag{44.80}$$

Como $\varepsilon > 0$ es arbitrario, esto significa que $\lim_{\delta_\tau \to 0} s_\tau = I_*$.

La desigualdad $s_\tau \leqslant I_*$ proviene directamente de la definición de integral inferior I_* (véase (44.76)). Por eso basta sólo demostrar la desigualdad

$$s_\tau > I_* - \varepsilon \tag{44.81}$$

a condición de que $\delta_\tau < 10^{-k}$.

Supongamos que $S_k = S_k(X_0)$ se compone de los cubos $Q_1, \ldots, Q_m$. Por analogía con lo que se hizo al demostrar el teorema 10, designemos mediante P_j un cubo que se obtiene de Q_j por transformación de semejanza de centro en el centro

del cubo Q_j y la razón de semejanza igual a 3, $j = 1, 2, \ldots, m$. Pongamos

$$P = \bigcup_{j=1}^{m} P_j, \quad G = X \setminus P. \tag{44.82}$$

De las definiciones de los conjuntos P y G se desprende que el conjunto G está separado del poliedro $S_k(X_0)$ por una "franja" de cubos con las aristas de longitud 10^{-k}. Estimemos ante todo la medida μP. De la definición del conjunto P (véase (44.82)) y la desigualdad (44.79) tenemos (compárese con (44.63))

$$\mu P = \mu \bigcup_{j=1}^{m} P_j \leqslant \sum_{j=1}^{m} \mu P_j = 3^n \sum_{j=1}^{m} \mu Q_j = 3^n \mu S_k(X_0) < \frac{\varepsilon}{3M}. \tag{44.83}$$

Hemos de notar a continuación que para todo conjunto $A \subset X$ de diámetro $d(A) < 10^{-k}$, que se interseca con el conjunto $G : A \cap G \neq \varnothing$, existe y, además, el único conjunto $X_i^* \in \tau^*$ tal que

$$A \subset X_i^*. \tag{44.84}$$

En efecto, elijamos un punto $x \in A \cap G$. Puesto que $A \subset X$, entonces $x \in X$, y, por ende, el punto x está contenido en cierto elemento X_i^* de la partición τ^*. Para este elemento se cumple precisamente la inclusión (44.84). Efectivamente, si la inclusión dada no tuviera lugar, existiría un punto $y \in A \setminus X_i^*$. Ya que $x \in A$, $y \in A$ y $d(A) < 10^{-k}$, resulta que $\rho(x, y) < 10^{-k}$. Por consiguiente, el segmento que tiene por sus extremos los puntos x e y, la longitud inferior a 10^{-k} y el extremo x dispuesto en el conjunto G, no se interseca con el conjunto $S_k(X_0)$, pues está separado de G por una franja de ancho 10^{-k}. No obstante, de lo que un extremo del segmento pertenece a cierto conjunto, en el caso dado al conjunto X_i^*, y el otro no le pertenece, se infiere (véase el lema 9 en el punto 18.2) que en dicho segmento existe un punto $z \in \partial X_i^*$. Mas (véase (44.78)) $\partial X_i^* \subset X_0 \subset S_k(X_0)$, es decir, $z \in S_k(X_0)$. Por lo tanto, el segmento citado se interseca con el conjunto $S_k(X_0)$. La contradicción obtenida demuestra la inclusión (44.84).

Demostremos la unicidad del conjunto X_i^* que satisface la inclusión (44.84). Supongamos que existe un conjunto más, $X_k^* \in \tau^*$, tal que $A \subset X_k^*$, $k \neq i$. En este caso $A \subset X_i^* \cap X_k^*$. Si la intersección $X_i^* \cap X_k^*$ contuviera por lo menos un punto que fuera al mismo tiempo interior para los conjuntos X_i^* y X_k^*, este punto sería interior también para la intersección $X_i^* \cap X_k^*$ y entonces tendría lugar la desigualdad $\mu(X_i^* \cap X_k^*) > 0$. Esta desigualdad contradice la definición de partición (véase el p. 44.3), en virtud de la cual $\mu(X_i^* \cap X_k^*) = 0$ cuando $i \neq k$. Por consiguiente, todo punto de la intersección $X_i^* \cap X_k^*$, y, por ende, todo punto del conjunto A es un punto de frontera por lo menos para uno de los conjuntos X_i^*, X_k^*. Pero, en tal caso $A \subset \bigcup_{i=1}^{i_0} \partial X_i^* = X_0 \subset S_k(X_0)$. Esto es imposible, puesto que el conjunto A se interseca con el conjunto G el cual no se corta con $S_k(X_0)$. La contradicción se ha obtenido de la suposición sobre la existencia del segundo elemento X_k^* de τ^* que contiene el conjunto A. Por consiguiente, tal elemento es único.

Tomemos ahora una partición arbitraria $\tau = \{X_j\}_{j=1}^{j_0}$ del conjunto X de finura $\delta_\tau < 10^{-k}$. Dividamos la suma inferior de Darboux

$$s_\tau = \sum_{j=1}^{j_0} m_j \mu X_j,\ m_j = \inf_{x \in X_j} f(x),\ j = 1, 2, \ldots, j_0,$$

en dos sumandos correspondientes a aquellos X_j que se intersecan con el conjunto G y aquellos que no se intersecan con él y, por lo tanto, se disponen íntegramente en el conjunto P (véase (44.82)):

$$s_\tau = \sum_{X_j \cap G \neq \varnothing} m_j \mu X_j + \sum_{X_j \subset P} m_j \mu X_j. \tag{44.85}$$

Al emplear la desigualdad evidente

$$|m_j| \leqslant M, j = 1, 2, \ldots, j_0. \tag{44.86}$$

donde $|f(x)| \leqslant M$, $x \in X$, y la estimación (44.83), obtendremos

$$\left| \sum_{X_j \subset P} m_j \mu X_j \right| \leqslant \sum_{X_j \subset P} |m_j| \mu X_j \leqslant$$

$$\leqslant M \sum_{X_j \subset P} \mu X_j \leqslant M\mu \bigcup_{X_j \subset P} X_j \leqslant M\mu P < M \frac{\varepsilon}{3M} = \frac{\varepsilon}{3}.$$

En particular, $\sum_{X_j \subset P} m_j \mu X_j > -\frac{\varepsilon}{3}$. Por esto de (44.85) tenemos

$$s_\tau > \sum_{X_j \cap G \neq \varnothing} m_j \mu X_j - \frac{\varepsilon}{3}. \tag{44.87}$$

Diremos ahora que $d(X_j) \leqslant \delta_\tau < 10^{-k}$, razón por la cual para todo X_j, que se interseca con el conjunto G, existe, en virtud de (44.84) tal $X_i^* \in \tau^*$, que $X_j \subset X_i^*$. Designaremos mediante G_i la reunión de todos aquellos X_j que se intersecan con G y están contenidos en X_i^*:

$$G_i = \bigcup_{X_j \subset X_i^*,\ X_j \cap G \neq \varnothing} X_j.$$

Agrupando en la suma $\sum_{X_j \cap G \neq \varnothing} m_j \mu X_j$ los sumandos contenidos en un mismo conjunto G_i, escribamos la suma citada en la forma

$$\sum_{X_j \cap G \neq \varnothing} m_j \mu X_j = \sum_{i=1}^{i_0} \sum_{X_j \subset G_i} m_j \mu X_j. \tag{44.88}$$

Con el fin de estimar la suma interior, observemos que para todo $i = 1, 2, \ldots, i_0$, de acuerdo con la igualdad obvia

$$X_i^* = (X_i^* \cap G_i) \cup (X_i^* \setminus G_i) = G_i \cup (X_i^* \setminus G_i)$$

(la segunda igualdad proviene de la inclusión $G_i \subset X_i^*$), tenemos

$$m_i^* \mu X_i^* = m_i^* \mu G_i + m_i^* \mu (X_i^* \setminus G_i) =$$
$$= m_i^* \mu \bigcup_{X_j \subset G_i} X_j + m_i^* \mu (X_i^* \setminus G_i) =$$
$$= m_i^* \sum_{X_j \subset G_i} \mu X_j + m_i^* \mu (X_i^* \setminus G_i). \quad (44.89)$$

Estimemos el segundo sumando. Todo punto $x \in X_i^* \setminus G_i$ pertenece a cierto conjunto $X_j \in \tau : x \in X_j$. Este conjunto X_j no puede intersecarse con G, puesto que todo $X_j \in \tau$, que se interseca con G, está contenido íntegramente en cierto elemento de la partición τ^* (véase (44.84)). Por cuanto la intersección $X_j \cap X_i^*$ no es vacía: $x \in X_j \cap X_i^*$, a título de este elemento en el caso dado puede intervenir sólo el conjunto X_i^*, es decir, $X_j \subset X_i^*$. Mas, en vista de la definición del conjunto G_i, tendría lugar en estas circunstancias la inclusión $X_j \subset G_i$, y, por lo tanto, $x \in G_i$. Esto contradice la suposición de que $x \in X_i^* \setminus G_i$. Así pues, el conjunto X_j no se interseca con G y, por ende, $X_j \subset P$. De aquí se deduce, en particular, que $x \in P$. Ya que x es un punto arbitrario del conjunto $X_i^* \setminus G_i$, se tiene $X_i^* \setminus G_i \subset P$, por lo cual $X_i^* \setminus G_i \subset X_i^* \cap P$.

Al hacer uso de dicha inclusión y de la desigualdad (44.86), tenemos

$$m_i^* \mu (X_i^* \setminus G_i) \leqslant M \mu (X_i^* \cap P).$$

Sustituyendo esta desigualdad en (44.89), tendremos

$$m_i^* \mu X_i^* \leqslant \sum_{X_j \subset G_i} m_j^* \mu X_j + M \mu (X_i^* \cap P).$$

Ha de notarse ahora que de la inclusión $X_j \subset G_i \subset X_i^*$ proviene la desigualdad $m_i^* \leqslant m_j$ (la cota inferior de un subconjunto no es menor que la cota inferior del propio conjunto) y obtenemos

$$m_i^* \mu X_i^* \leqslant \sum_{X_j \subset G_i} m_j \mu X_j + M \mu (X_i^* \cap P),$$

de donde

$$\sum_{X_j \subset G_i} m_j \mu X_j \geqslant m_i^* \mu X_i^* - M \mu (X_i^* \cap P).$$

Al sumar ambos miembros según i desde 1 hasta i_0, en virtud de (44.88), tendremos

$$\sum_{X_j \cap G \neq \varnothing} m_j \mu X_j \geqslant \sum_{i=1}^{i_0} m_i^* \mu X_i - M \sum_{i=1}^{i_0} \mu (X_i^* \cap P) = s_{\tau^*} -$$
$$- M \sum_{i=1}^{i_0} \mu (X_i^* \cap P).$$

Puesto que, de acuerdo con (44.83),

$$\sum_{i=1}^{i_0} \mu(X_i^* \cap P) = \mu \bigcup_{i=1}^{i_0} X_i^* \cap P \leqslant \mu P < \frac{\varepsilon}{3M},$$

resulta

$$\sum_{X_j \cap G \neq \varnothing} m_j \mu X_j > s_{\tau^*} - \frac{\varepsilon}{3}. \tag{44.90}$$

Al aplicar, ahora, sucesivamente las desigualdades (44.87), (44.90) y (44.77), obtendremos

$$s_\tau > \sum_{X_j \cap G \neq \varnothing} m_j \mu X_j - \frac{\varepsilon}{3} > s_{\tau^*} - \frac{2\varepsilon}{3} > I_* - \varepsilon,$$

es decir, la desigualdad (44.81) y, por lo tanto, el teorema 11 se han demostrado. □

Con ayuda de dicho teorema pueden enunciarse dos criterios de integrabilidad de una función acotada.

Teorema 12 (criterio de Darboux). *Una función acotada en un conjunto medible según Jordan es integrable según Riemann cuando, y sólo cuando, sus integrales superior e inferior de Darboux son iguales.*

DEMOSTRACIÓN. Sean I_* e I^* las integrales de Darboux, inferior y superior, respectivamente, de la función f que es acotada en un conjunto medible X. Por consiguiente, para cualquier partición τ del conjunto X se verifican las desigualdades (véase (44.76))

$$s_\tau \leqslant I_* \leqslant I^* \leqslant S_\tau. \tag{44.91}$$

NECESIDAD DE LA CONDICIÓN $I_* = I^*$. Si la función f es integrable en el conjunto X, se tiene (véase (44.57))

$$\lim_{\delta_n \to 0} (S_\tau - s_\tau) = 0,$$

y dado que $0 \leqslant I^* - I_* \leqslant S_\tau - s_\tau$, entonces $I_* = I^*$.

SUFICIENCIA DE LA CONDICIÓN $I_* = I^*$. Si $I_* = I^*$, en virtud del teorema 11, tenemos

$$\lim_{\delta_\tau \to 0} (S_\tau - s_\tau) = \lim_{\delta_\tau \to 0} S_\tau - \lim_{\delta_\tau \to 0} s_\tau = I^* - I_* = 0,$$

y, en tal caso, de acuerdo con el teorema 8 del p. 44.4, la función f es integrable. □

Teorema 13 (criterio de Riemann). *Una función f, acotada en el conjunto medible según Jordan X, es integrable según Riemann cuando, y sólo cuando, para todo $\varepsilon > 0$ existe tal partición τ del conjunto X que*

$$S_\tau - s_\tau < \varepsilon, \tag{44.92}$$

donde s_τ y S_τ son las sumas de Darboux inferior y superior de la función f, correspondientes a la partición τ.

DEMOSTRACIÓN. Si la función f es integrable en el conjunto X, para ella se cumple la condición (44.57) (véase el teorema 8 en el p. 44.4). La validez de (44.92) se deduce de la definición de límite de las sumas de Darboux para $\delta_\tau \to 0$.

Si, en cambio, se cumple la condición (44.92), entonces, en virtud de (44.91), con cualquier $\varepsilon > 0$ es válida la desigualdad $0 \leqslant I^* - I_* < \varepsilon$, y, por eso, $I_* = I^*$. De aquí, de conformidad con el teorema 12, se desprende que la función f es integrable en el conjunto X. □

Así pues, recordando la definición de la integral múltiple, aducida en el p. 44.3, el teorema 8 del p. 44.4 y los teoremas 12 y 13 del punto presente, llegamos a que las cinco afirmaciones que siguen son equivalentes:

1) *la función f es integrable en el conjunto X, es decir, existe el límite* $\lim\limits_{\delta_\tau \to 0} \sigma_\tau = \int f(x)dX$;

2) $\lim\limits_{\delta_\tau \to 0} (S_\tau - s_\tau) = 0$;

3) $\lim\limits_{\delta_\tau \to 0} \sum\limits_{i=1}^{i_0} \omega(f; X_i)\mu X_i = 0, \tau = \{X_i\}_{i=1}^{i_0}$ *es la partición del conjunto X;*

4) *para cualquier $\varepsilon > 0$ existe tal partición del conjunto X que $S_\tau - s_\tau < \varepsilon$;*

5) $I_* = I^*$.

De este modo, el cumplimiento de cada una de estas condiciones es equivalente a la existencia de la integral $\int f(x)dX$, con la particularidad de que

$$\int f(x)dX = \lim_{\delta_\tau \to 0} \sigma_\tau = \lim_{\delta_\tau \to 0} s_\tau = \lim_{\delta_\tau \to 0} S_\tau.$$

OBSERVACIÓN 1. Los teoremas obtenidos permiten ahora demostrar sin dificultad la aditividad de una integral extendida a los conjuntos medibles para funciones acotadas (véase el p. 44.6) en la forma siguiente: *si una función acotada f es integrable en los conjuntos disjuntos X_1 y X_2, es integrable también en el conjunto $X = X_1 \cup X_2$.*

En efecto si la función f está acotada y es integrable en los conjuntos X_1 y X_2, entonces, en virtud del teorema 13, para todo $\varepsilon > 0$ existen las particiones τ_1 y τ_2 de los conjuntos respectivos X_1 y X_2 de tal índole que

$$S_{\tau_1} - s_{\tau_1} < \frac{\varepsilon}{2}, \quad S_{\tau_2} - s_{\tau_2} < \frac{\varepsilon}{2}. \tag{44.93}$$

Como $\tau = \tau_1 \cup \tau_2$ es una partición del conjunto $X = X_1 \cup X_2$ y las sumas de Darboux, superior S_τ e inferior s_τ, que corresponden a esta partición, se expresan en términos de las sumas análogas de Darboux correspondientes a la particiones τ_1 y τ_2 según las fórmulas $S_\tau = S_{\tau_1} + S_{\tau_2}$, $s_\tau = s_{\tau_1} + s_{\tau_2}$, entonces, restando de la primera de las igualdades dadas la segunda, obtenemos, en vista de (44.93)

$$S_\tau - s_\tau = (S_{\tau_1} - s_{\tau_1}) + (S_{\tau_2} - s_{\tau_2}) < \varepsilon.$$

Del cumplimiento de esta condición se deduce (también conforme al teorema 13) que la función f es integrable en el conjunto X.

OBSERVACIÓN 2. Como ya se ha dicho en el p. 44.3, para las funciones de una variable, definidas en los segmentos, disponemos de dos definiciones de la integral, a

saber, una, dada en el p. 27.1, con ayuda de las particiones de los segmentos en segmentos y sólo segmentos, y la otra definición del p. 44.3, con ayuda de las particiones de los segmentos en cualesquiera conjuntos medibles según Jordan. Estas dos definiciones son equivalentes.

Demostrémoslo. Para ambas definiciones la condición necesaria de intégrabilidad consiste en el carácter acotado de la función en consideración: véase el teorema 1 en el p. 27.2 y la observación al teorema 7 en el p. 44.4 (un segmento es la clausura del intervalo, es decir, la clausura de un conjunto abierto). Por ello consideraremos la función f que está acotada en cierto segmento $[a, b]$. Supongamos que para esta función existe la integral $I = \lim_{\delta_\tau \to 0} \sum_{i=1}^{i_0} f(\xi_i)\mu X_i$ en el sentido del p. 44.3, es decir, para cualesquiera particiones $\tau = \{X_i\}_{i=1}^{i_0}$ del segmento $[a, b]$ en conjuntos X_i medibles según Jordan. En este caso si nos limitamos sólo a una parte de particiones τ, para las cuales todos los conjuntos X_i son segmentos, entonces para $\delta_\tau \to 0$, el límite de las sumas integrales $\sum_{i=1}^{i_0} f(\xi_i)\mu X_i$ referente a la parte mencionada de particiones existirá también y será igual al mismo número I. Por consiguiente, si existe una integral en el sentido del p. 44.3, existe la misma también en el sentido del p. 27.1.

Viceversa, supongamos que existe la integral $I = \int_a^b f(x)dx$ en el sentido del p. 27.1. Entonces, de acuerdo con el teorema 2 del p. 27.4, $\lim_{\delta_\tau \to 0} (S_\tau - s_\tau) = 0$, donde τ es la partición del segmento $[a, b]$ en segmentos. Por consiguiente, para todo $\varepsilon > 0$ existe tal $\delta > 0$ que para cualquier partición τ del segmento $[a, b]$ en segmentos cuyas longitudes no son superiores a δ, resulta lícita la desigualdad $S_\tau - s_\tau < \varepsilon$. Pero ya del único hecho consistente en la existencia por lo menos de una partición τ, para la cual se cumple la desigualdad $S_\tau - s_\tau < \varepsilon$, proviene, de conformidad con el teorema 13 del presente punto, que la función f es integrable en el sentido del p. 44.3.

Así pues, ambas definiciones de la integral referida a un segmento son realmente equivalentes.

OBSERVACIÓN 3. De lo demostrado se desprende también el siguiente reforzamiento de las condiciones suficientes de integrabilidad de las funciones, demostradas en el teorema 2 del p. 24.4: para que una función f sea integrable en el segmento $[a, b]$ en el sentido de la definición de integral en el p. 27.1, es suficiente que para todo $\varepsilon > 0$ exista por lo menos tal partición τ del segmento $[a, b]$ en segmentos que para las sumas inferiores y superiores de Darboux, correspondientes a dicha partición, se verifique la desigualdad $S_\tau - s_\tau < \varepsilon$.

En efecto, en tal caso la función f es integrable en el segmento $[a, b]$ en el sentido del p. 44.3, por lo cual, de acuerdo con lo demostrado, también en el sentido del p. 27.1.

OBSERVACIÓN 4. De la observación antecedente se infiere directamente que la función f, acotada en cierto segmento $[a, b]$ e integrable según Riemann en cual-

quier segmento $[a, \eta]$, $a < \eta < b$, es integrable también en todo el segmento $[a, b]$ (este hecho se ha indicado sin demostración en el p. 33.1). Efectivamente, si $|f(x)| \leqslant M$, $x \in [a, b]$, y se ha dado $\varepsilon > 0$, entonces elijamos δ, $0 < \delta < b - a$, de modo tal que sea $\delta < \dfrac{\varepsilon}{4M}$. Por ser la función f integrable en el segmento $[a, b - \delta]$, existe tal partición τ de éste, que si s_τ y S_τ son las sumas inferior y superior de Darboux de la función f para esta partición, se tiene

$$S_\tau - s_\tau < \frac{\varepsilon}{2}.$$

Designemos con τ_0 la partición del segmento $[a, b]$ que se obtiene de la partición τ_0 del segmento $[a, b - \delta]$ por adición del punto b: $\tau_0 = \tau \cup \{b\}$ y sea $m_0 = \inf\limits_{[b-\delta, b]} f(x)$, $M_0 = \sup\limits_{[b-\delta, b]} f(x)$. Si s_{τ_0} y S_{τ_0} son las sumas inferior y superior de Darboux de la función f para la partición τ_0, entonces

$$S_{\tau_0} = S_\tau + M_0\delta, \quad s_{\tau_0} = s_\tau + m_0\delta.$$

Por esta razón

$$S_{\tau_0} - s_{\tau_0} = S_\tau - s_\tau + (M_0 - m_0)\delta < \frac{\varepsilon}{2} + 2M\delta = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

y, por lo tanto, de acuerdo con la observación 3, la función f es integrable según Riemann en el segmento $[a, b]$.

§ 45. REDUCCIÓN DE LA INTEGRAL MÚLTIPLE A UNA REITERADA

Pasemos ahora a las propiedades de la integral múltiple relacionadas con los rasgos específicos que distinguen el caso multidimensional del unidimensional. El empleo de estas propiedades facilita considerablemente con frecuencia el cálculo de las integrales múltiples concretas. Las demostraciones completas se darán sólo para el caso de las funciones de dos variables. El caso general n-dimensional no difiere, teóricamente, del caso plano, no obstante los razonamientos en el caso n-dimensional adoptan una forma más engorrosa y, por lo tanto, difícilmente contemplada.

45.1. REDUCCIÓN DE LA INTEGRAL DOBLE A UNA REITERADA

En el presente párrafo se mostrará que la integración de las funciones de varias variables puede reducirse a la integración sucesiva de las funciones de una sola variable. Comenzaremos por definir el concepto de integral reiterada.

Supongamos que en el segmento $[a, b]$ se han dado las funciones $\varphi(x)$ y $\psi(x)$ tales que $\varphi(x) \leqslant \psi(x)$, $a \leqslant x \leqslant b$, y que en el conjunto (fig. 185)

$$X = \{(x, y) : a \leqslant x \leqslant b, \quad \varphi(x) \leqslant y \leqslant \psi(x)\} \tag{45.1}$$

está definida la función $f(x, y)$.

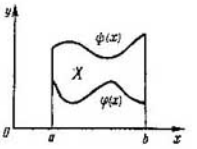


Fig. 185

Si la función $f(x, y)$, como función de la variable y, es integrable en el segmento $[\varphi(x), \psi(x)]$ para todo $x \in [a, b]$ fijo, es decir, si para todo $x \in [a, b]$ existe la integral $\int_{\varphi(x)}^{\psi(x)} f(x, y)dx$ y la función

$$F(x) = \int_{\varphi(x)}^{\psi(x)} f(x, y)dy \tag{45.2}$$

es integrable en el segmento $[a, b]$, entonces la integral

$$\int_a^b \left[\int_{\varphi(x)}^{\psi(x)} f(x, y)dy \right] dx \tag{45.3}$$

se llama *reiterada* y se designa mediante

$$\int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y)\, dy. \tag{45.4}$$

La función $F(x)$, definida por la igualdad (45.2), se denomina *integral dependiente del parámetro* x. De este modo, la integral reiterada (45.4) es una integral de la integral dependiente de parámetro (véanse también §§ 53, 54).

Observemos que el conjunto X, que se define por la fórmula (45.1), es medible en el sentido de la medida plana de Jordan y cerrado. En efecto, de la continuidad de las funciones φ y ψ en el segmento $[a, b]$ proviene su carácter acotado, por lo cual el conjunto X está acotado. Luego, su frontera ∂X se compone de las gráficas de las funciones citadas φ, ψ y también, quizás, de los segmentos de las rectas $x = a$ y $x = b$. Cada uno de los conjuntos mencionados tiene medida cero (véase el teorema 3 en el p. 44.2), razón por la cual la frontera ∂X del conjunto X tiene también medida cero. En fin, el conjunto X se da con ayuda de las desigualdades no estrictas $a \leqslant x \leqslant b$, $\varphi(x) \leqslant y \leqslant \psi(x)$, donde las funciones φ y ψ son continuas, a consecuencia de lo cual dichas desigualdades quedan vigentes también al pasar al límite, de lo que precisamente proviene el carácter cerrado del conjunto X. De este modo, X es un compacto medible.

Las condiciones suficientes para que haya la posibilidad de reducir la integral doble a una reiterada se proporcionan por el siguiente teorema.

Teorema 1. *Sea $f(x, y)$ una función continua en el conjunto X dado mediante la fórmula* (45.1). *En este caso*

$$\iint_X f(x, y)\,dx\,dy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y)\,dy. \tag{45.5}$$

A la demostración del teorema antepongamos el lema siguiente.

Lema 1. *En las suposiciones del teorema 1 la función* (45.2) *es continua en el segmento* $[a, b]$.

DEMOSTRACIÓN DEL LEMA. Señalemos ante todo que la integral (45.2) existe para todo $x \in [a, b]$. Efectivamente, la función $f(x, y)$, siendo continua por totalidad de las variables x e y, es continua respecto de cada una de ellas. Por esto, la integral citada existe como integral de una función continua respecto de y en el segmento $[\varphi(x), \psi(x)]$.

Al sustituir en esta integral la variable y por t, rigiéndose por la fórmula

$$y = \varphi(x) + [\psi(x) - \varphi(x)]\,t,\ 0 \leqslant t \leqslant 1, \tag{45.6}$$

obtendremos

$$F(x) = \int_0^1 f[x, \varphi(x) + (\psi(x) - \varphi(x))t]\,(\psi(x) - \varphi(x))\,dt. \tag{45.7}$$

Pongamos

$$g(x, t) = f[x, \varphi(x) + (\psi(x) - \varphi(x))t]\,(\psi(x) - \varphi(x)).$$

Como la función $g(x, t)$ se obtiene con ayuda de una operación aritmética y una composición de las funciones continuas f, φ, ψ y (45.6), entonces, en virtud del teorema sobre funciones continuas (véase el p. 19.5), $g(x, t)$ es continua por totalidad de las variables x, t en el rectángulo

$$\Delta = [(x, t) : a \leqslant x \leqslant b,\ 0 \leqslant t \leqslant 1].$$

De este modo, para la función $F(x)$ (véase (45.2)) tiene lugar, debido a (45.7), una representación más simple

$$F(x) = \int_0^1 g(x, t)\,dt$$

(más simple porque los límites de integración en ella son constantes).

Sea ahora $x \in [a, b]$, $x + \Delta x \in [a, b]$.

Designemos mediante $\omega(\delta; g)$ el módulo de continuidad (véase el p. 19.7) de la función $g(x, t)$. En este caso

$$|F(x + \Delta x) - F(x)| = \left| \int_0^1 g(x + \Delta x, t)\,dt - \int_0^1 g(x, t)\,dt \right| \leqslant$$

$$\leqslant \int_0^1 |g(x + \Delta x, t) - g(x, t)|\,dt \leqslant \omega(|\Delta x|; g). \tag{45.8}$$

La función $g(x, t)$, siendo continua en el conjunto acotado cerrado Δ, es en éste uniformemente continua, por lo cual (véase el p. 19.7), $\lim_{\delta \to 0} \omega(\delta; g) = 0$. De aquí, en virtud de la desigualdad (45.8), tenemos:

$$\lim_{\Delta x \to 0} [F(x + \Delta x) - F(x)] = 0,$$

lo que precisamente implica la continuidad de la función $F(x)$, definida por la fórmula (45.2). □

DEMOSTRACIÓN DEL TEOREMA. Ha de notarse ante todo que la integral en el segundo miembro de la igualdad (45.5), es decir,

$$\int_a^b F(x)dx = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y)\, dy$$

es la integral de una función continua (véase el lema) y por esta razón existe.

Partiremos ahora el conjunto X en las partes X_{ij}, $i, j = 1, 2, \ldots, k$, de la manera siguiente. Consideraremos la partición $\tau_k = \{x_i\}_{i=0}^{i=k}$ del segmento $[a, b]$ en k segmentos iguales:

$$a = x_0 < x_1 < \ldots < x_k = b,$$

$$x_i - x_{i-1} = \frac{b - a}{k}, \; i = 1, 2, \ldots, k,$$

y sea

$$\varphi_0(x) = \varphi(x),$$

$$\varphi_1(x) = \varphi(x) + \frac{1}{k}[\psi(x) - \varphi(x)],$$

.....

$$\varphi_j(x) = \varphi(x) + \frac{j}{k}[\psi(x) - \varphi(x)],$$

.....

$$\varphi_k(x) = \varphi(x) + \frac{k}{k}[\psi(x) - \varphi(x)] = \psi(x).$$

Pongamos $X_{ij} = \{(x, y) : x_{i-1} \leqslant x \leqslant x_i, \varphi_{j-1}(x) \leqslant y \leqslant \varphi_j(x)\}$, y sea $\tau_k^* = \{X_{ij}\}$, $i, j = 1, 2, \ldots, k$. Evidentemente, τ_k^* es la partición del conjunto X (fig. 186). Ahora tenemos

$$\int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y)dy = \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} dx \int_{\varphi(x)}^{\psi(x)} f(x, y)dy =$$

$$= \sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} dx \sum_{j=1}^{k} \int_{\varphi_{j-1}(x)}^{\varphi_j(x)} f(x, y)dy =$$

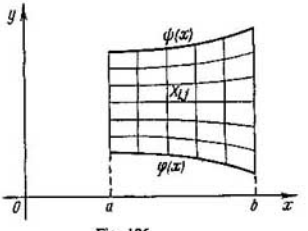


Fig. 186

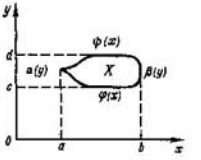


Fig. 187

$$= \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{x_{i-1}}^{x_i} dx \int_{\varphi_{j-1}(x)}^{\varphi_j(x)} f(x, y)dy. \quad (45.9)$$

Pongamos

$$m_{ij} = \inf_{X_{ij}} f(x, y) \text{ y } M_{ij} = \sup_{X_{ij}} f(x, y), \; i, j = 1, 2, \ldots, k.$$

Observando que

$$\mu X_{ij} = \int_{x_{i-1}}^{x_i} [\varphi_j(x) - \varphi_{j-1}(x)]\, dx,$$

obtendremos

$$\int_{x_{i-1}}^{x_i} dx \int_{\varphi_{j-1}(x)}^{\varphi_j(x)} f(x, y)dy \leqslant M_{ij} \int_{x_{i-1}}^{x_i} dx \int_{\varphi_{j-1}(x)}^{\varphi_j(x)} dy =$$

$$= M_{ij} \int_{x_{i-1}}^{x_i} [\varphi_j(x) - \varphi_{j-1}(x)]\, dx = M_{ij}\, \mu X_{ij}. \quad (45.10)$$

y, por analogía,

$$\int_{x_{i-1}}^{x_i} dx \int_{\varphi_{j-1}(x)}^{\varphi_j(x)} f(x, y)dy \geqslant m_{ij}\, \mu X_{ij}. \quad (45.11)$$

Con ayuda de las desigualdades (45.10) y (45.11) para la integral reiterada (45.9) obtenemos la siguiente estimación en términos de las sumas inferiores y superiores de Darboux s_{τ_k} y S_{τ_k} de la función $f(x, y)$:

$$s_{\tau_k} = \sum_{i=1}^{k} \sum_{j=1}^{k} m_{ij}\, \mu X_{ij} \leqslant \int_a^b dx \int_{\psi(x)}^{\Psi(x)} f(x, y)dy \leqslant$$

$$\leqslant \sum_{i=1}^{k} \sum_{j=1}^{k} M_{ij}\, \mu X_{ij} = S_{\tau_k}. \quad (45.12)$$

Para la finura $\delta_{\tau_k^*}$ de la partición τ_k^* de la región G tenemos $\lim_{k\to\infty} \delta_{\tau_k^*} = 0$. En efecto, como ya se ha indicado, las funciones φ y ψ están acotadas en el segmento $[a, b]$ debido a su continuidad, es decir, existe tal constante $M > 0$ que $|\varphi(x)| \leqslant M$ y $|\psi(x)| \leqslant M$, cualesquiera que sean $x \in [a, b]$. Por ello, para el diámetro $d(X_{ij})$ de cada conjunto $X_{ij} \in \tau_k^*$ tenemos

$$dX_{ij} \leqslant \sqrt{\left(\frac{b-a}{k}\right)^2 + [\max_{\substack{x' \in [x_{i-1}, x_i] \\ x'' \in [x_{i-1}, x_i]}} |\varphi_j(x'') - \varphi_{j-1}(x')|]^2} =$$

$$= \sqrt{\left(\frac{b-a}{k}\right)^2 + \max_{\substack{x' \in [x_{i-1}, x_i] \\ x'' \in [x_{i-1}, x_i]}} |[\varphi_j(x'') - \varphi_j(x')] + [\varphi_j(x') - \varphi_{j-1}(x')]|^2}.$$

Observando que para $x' \in [x_{i-1}, x_i]$, $x'' \in [x_{j-1}, x_i]$, en virtud de la definición de las funciones φ_j, se cumplen las desigualdades

$$|\varphi_j(x'') - \varphi_j(x')| = |\varphi(x'') - \varphi(x')| + \frac{j}{k}|\varphi(x'') - \varphi(x')| +$$

$$+ \frac{j}{k}|\psi(x'') - \psi(x')| \leqslant 2\omega\left(\frac{b-a}{k}, \varphi\right) + \omega\left(\frac{b-a}{k}, \psi\right),$$

$$j = 0, 1, \ldots, k,$$

donde $\omega(\delta, \varphi)$ y $\omega(\delta, \psi)$ son los módulos de continuidad de las funciones φ y ψ, respectivamente, así como para todos $x \in [a, b]$ las desigualdades

$$|\varphi_j(x) - \varphi_{j-1}(x)| = \left|\frac{j}{k}[\psi(x) - \varphi(x)] - \frac{j-1}{k}[\psi(x) - \varphi(x)]\right| =$$

$$= \frac{|\psi(x) - \varphi(x)|}{k} \leqslant \frac{|\psi(x)| + |\varphi(x)|}{k} \leqslant \frac{2M}{k}, \quad j = 1, 2, \ldots, k,$$

obtendremos

$$d(X_{ij}) \leqslant \sqrt{\left(\frac{b-a}{k}\right)^2 + \max_{[x_{j-1}, x_i]} [\varphi_j(x'') - \varphi_{j-1}(x')]^2} \leqslant$$

$$\leqslant \sqrt{\left(\frac{b-a}{k}\right)^2 + \left[\omega\left(\frac{b-a}{k}, \psi\right) + 2\omega\left(\frac{b-a}{k}, \varphi\right) + \frac{2M}{k}\right]^2}.$$

Por consiguiente, $\delta_{\tau_k^*} = \max_{i,j} d(X_{ij}) \to 0$ cuando $k \to +\infty$. Por esta razón, siendo la función $f(x, y)$ integrable en X (véase el p. 44.4), tenemos

$$\lim_{k\to\infty} s_{\tau_k^*} = \lim_{k\to\infty} S_{\tau_k^*} = \iint_G f(x, y)\,dx\,dy.$$

Pasando ahora al límite en la desigualdad (45.12) para $k \to \infty$, obtendremos la fórmula (45.5). □

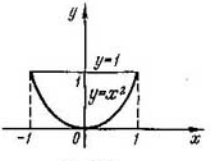


Fig. 188

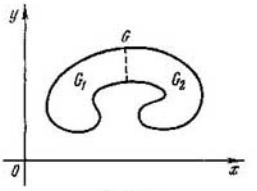


Fig. 189

Si el conjunto X es de tal género que existen las funciones continuas $\alpha(y)$ y $\beta(y)$, $\alpha(y) \leqslant \beta(y)$, $c \leqslant y \leqslant d$ tales que

$$X = \{(x, y) : c \leqslant y \leqslant d, \alpha(y) \leqslant x \leqslant \beta(y)\}, \tag{45.13}$$

mientras que la función $f(x, y)$ es, como hasta ahora, continua en $\bar{X}$, entonces, en virtud de que la variable x no tiene ventaja alguna ante la variable y, del teorema 1 proviene que

$$\iint_X f(x,y)dx\,dy = \int_c^d dy \int_{\alpha(y)}^{\beta(y)} f(x,y)dx. \tag{45.14}$$

En cambio, si para el conjunto X son justas tanto la igualdad (45.1), como también (45.13) (fig. 187), entonces igualando entre sí los segundos miembros de las igualdades (45.5) y (45.14), para la función $f(x, y)$, continua en el conjunto $\bar{X}$, obtendremos la fórmula

$$\int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x,y)dy = \int_c^d dy \int_{\alpha(y)}^{\beta(y)} f(x,y)dx, \tag{45.15}$$

que expresa la regla de cambio del orden de integración en las integrales reiteradas.

Se debe decir que las condiciones, bajo las cuales se han demostrado las fórmulas (45.5), (45.14) y (45.15), pueden ser debilitadas.

Ejemplo. Calculemos la integral de la función $z = x^2y$, extendida a la región finita G que está limitada por una parte de la parábola $y = x^2$ y la recta $y = 1$ (fig. 188). Tenemos

$$\iint_G x^2y\,dx\,dy = \int_{-1}^{1} x^2dx \int_{x^2}^{1} y\,dy =$$

$$= \int_{-1}^{1} \frac{y^2}{2}\Big|_{x^2}^{1} x^2dx = \frac{1}{2}\int_{-1}^{1}(1 - x^4)x^2dx = \frac{4}{21}.$$

Si se requiere calcular una integral doble referida a un conjunto que no puede darse en la forma (45.1) ó (45.13), entonces con el fin de emplear las fórmulas obtenidas se debe tratar de dividir el conjunto dado en unas partes, cada una de las cuales ya tendrá la forma (45.1) ó (45.13) (fig. 189). Si se logra alcanzarlo, entonces, por ser aditiva la integral en los conjuntos (véase el p. 44.6), el cálculo de la integral dada se reducirá al cálculo de las integrales extendidas a las partes mencionadas; las

últimas integrales pueden ser reducidas, con ayuda de las fórmulas (45.5) y (45.14), a las de multiplicidad unitaria.

Ejercicios. Calcúlense las integrales:

1. $\iint\limits_X \frac{dx\,dy}{y}$, $X = \{(x, y) : x^2 - 6x - 5 < 0; y > 1; 3x - y - 2 > 0, x^2 - y > 0\}$.

2. $\iint\limits_X x^2y^2dx\,dy$, $X = \{(x, y) : y > 0; xy < 1; x^2 - 3xy + 2y^2 < 0\}$.

3. $\iint\limits_X x\,dx\,dy$, $X = \{(x, y) : x < 20; y < 20; x - y + 5 > 0, xy > 6\}$.

4. $\iint\limits_X x\sqrt{1 + xy}\,dx\,dy$, $X = \left\{x, y) : xy < 1, x - 1 < \frac{xy}{x + 1}\right\}$.

5. $\int\limits_0^1 \int\limits_y^1 e^{x^2}dx\,dy$.

6. $\int\limits_0^{1/2} \int\limits_{2y}^1 \cos(x^2 + 1)dx\,dy$.

7. $\int\limits_0^1 \int\limits_{\sqrt{y}}^1 \operatorname{sen}(x^3 - 1)dx\,dy$.

8. $\int\limits_0^1 \int\limits_0^{1-y} e^{-x^2 + 2x}dx\,dy$.

9. $\int\limits_0^1 \int\limits_{\sqrt{y}}^1 e^{x^3/3}dx\,dy$.

10. $\int\limits_0^1 \int\limits_{(y-1)/2}^1 \operatorname{tg}(x^2 + x)dx\,dy$.

Simplifíquense las expresiones por cambio del orden de integración (la función f es continua en toda la región de integración):

11. $\int\limits_0^1 dy \int\limits_{y^2/9}^{\sqrt{x}} f(x, y)dx + \int\limits_1^3 dx \int\limits_{y^2/9}^1 f(x, y)dy$.

12. $\int\limits_0^1 dx \int\limits_0^{x^2} f(x, y)\,dy + \int\limits_1^4 dx \int\limits_0^{1/x} f(x, y)\,dy$.

13. $\int\limits_0^1 dx \int\limits_0^{(3-x)/2} f(x, y) + \int\limits_1^3 dx \int\limits_0^1 f(x, y)\,dy$.

14. Demuéstrese la fórmula de Dirichlet

$$\int\limits_a^b dx \int\limits_a^x f(x, y)dy = \int\limits_a^b dy \int\limits_y^b f(x, y)dx.$$

45.2. GENERALIZACIÓN PARA EL CASO n-DIMENSIONAL

Consideraremos primero el caso tridimensional. Supongamos que $X \subset R^3$ y la función $f(x, y, z)$ está definida en X. Designemos mediante X_{xy} la proyección del conjunto X sobre el plano coordenado de las variables x e y (fig. 190):

$$X_{xy} = \{(x, y, 0): \text{existe tal } z \text{ que } (x, y, z) \in X\}.$$

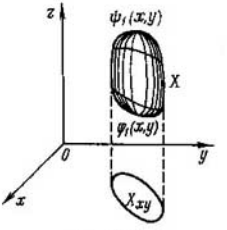


Fig. 190

Si el conjunto X es de la forma

$$X = \{(x, y, z) : (x, y) \in X_{xy},\ \varphi_1(x, y) \leqslant z \leqslant \psi_1(x, y)\},$$

donde las funciones $\varphi_1(x, y)$ y $\psi_1(x, y)$ son continuas en el conjunto X_{xy}, el cual es representable, a su vez, en la forma (45.1), y la función $f(x, y, z)$ es continua en el conjunto de partida X, entonces resulta válida la fórmula, análoga a la (45.5),

$$\iiint_X f(x, y, z)dxdydz = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} dy \int_{\varphi_1(x,y)}^{\psi_1(x,y)} f(x, y, z)dz. \tag{45.16}$$

Al reunir en el segundo miembro dos integrales exteriores, podemos escribir (45.16) en la forma

$$\iiint_X f(x, y, z)dxdydz = \iint_{X_{xy}} dxdy \int_{\varphi_1(x,y)}^{\psi_1(x,y)} f(x, y, z)dz. \tag{45.17}$$

Designemos ahora con $X(x)$ las secciones del conjunto X mediante los planos perpendiculares al eje de coordenadas Ox

$$X(x_0) = X \cap \{(x, y, z) : x = x_0\}.$$

Al reunir en el segundo miembro de la fórmula (45.16) dos integrales interiores, obtendremos:

$$\iiint_X f(x, y, z)dxdydz = \int_a^b dx \iint_{X(x)} f(x, y, z)dydz. \tag{45.18}$$

De este modo, las fórmulas (45.17) y (45.18) muestran que en el caso tridimensional existen dos métodos de reducción de la integral tridimensional a una reiterada que contiene integrales de menor multiplicidad.

En el caso particular, cuando $f(x, y, z) \equiv 1$, tenemos (véase la propiedad 1° de las integrales múltiples en el p. 44.6) $\iiint dxdydz = \mu X$, (μX es el volumen del conjunto X), $\iint_{X(x)} dydz = \mu X(x)$, ($\mu X(x)$ es el área de la sección $X(x)$).

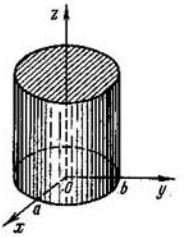


Fig. 191

De este modo

$$\mu X = \int_a^b \mu X(x)dx, \tag{45.19}$$

es decir, el volumen de un cuerpo es igual a la integral del área variable de la sección $X(x)$.

Ejemplo. Hallemos el volumen de un cilindro elíptico de altura h, de cuya base sirve una elipse de semiejes a y b. Al tomar como plano de coordenadas xy el plano de una de las bases del cilindro y como eje z, el eje de simetría del cilindro, perpendicular a las bases (fig. 191), obtendremos, de acuerdo con la fórmula (45.19), $\mu X = \int_0^h \mu X(z)dz$. Mas $X(z)$ es una elipse cuyos ejes son a y b, por lo cual (véase el ejemplo 4 en el p. 32.1) $\mu X(z) = \pi ab$, y, por consiguiente, $\mu X = \pi ab \int_0^h dz = \pi abh$.

Por analogía con el caso tridimensional, las integrales múltiples de una función de cualquier número de variables $n > 3$ pueden reducirse a las integrales reiteradas. Supongamos que R^n es un espacio n-dimensional, R^{n-1} un hiperplano $x_n = 0$, $X \subset R^n$ y $X_{x_1 \ldots x_{n-1}}$, la proyección del conjunto X sobre el hiperplano de las variables $x_1, \ldots, x_{n-1}$, es decir, sobre R^{n-1}.

$$X_{x_1 \ldots x_{n-1}} = \{(x_1, \ldots, x_{n-1}, 0): \text{existe tal } x_n \quad \text{que } (x_1, \ldots, x_{n-1}, x_n) \in X\}.$$

Supongamos, además, que existen tales funciones $\varphi(x_1, \ldots, x_{n-1})$ y $\psi(x_1, \ldots, x_{n-1})$, continuas en $X_{x_1 \ldots x_{n-1}}$, que el conjunto X se compone de los puntos $x = (x_1, \ldots, x_{n-1}, x_n)$, para los cuales

$$(x_1, \ldots, x_{n-1}) \in X_{x_1, \ldots, x_{n-1}}, \qquad \varphi(x_1, \ldots, x_{n-1}) \leqslant x_n \leqslant \psi(x_1, \ldots, x_{n-1}).$$

Sea el conjunto $X_{x_1 \ldots x_{n-1}}$ medible, en el sentido de la medida $(n-1)$-dimensional de Jordan, y cerrado. Entonces, por analogía con el caso bidimensional (véase el p. 45.1), se demuestra que X es también medible, pero ya en el sentido de la medida n-dimensional, y está cerrado, razón por la cual es un compacto.

Si la función $f(x) = f(x_1, \ldots, x_n)$ es continua en el compacto X, es lícita la fórmula

$$\overbrace{\int\cdots\int}^{n \text{ veces}}_{X} f(x_1, \ldots, x_n)dx_1 \ldots dx_n =$$

$$= \overbrace{\int\cdots\int}^{n-1 \text{ veces}}_{X_{x_1 \ldots x_{n-1}}} dx_1 \ldots dx_{n-1} \int_{\varphi(x_1, \ldots, x_{n-1})}^{\psi(x_1, \ldots, x_{n-1})} f(x_1, \ldots, x_{n-1})dx_n, \quad (45.20)$$

la cual reduce la integración de la función de n variables a la integración sucesiva de la función de una sola variable y de otra función de $n-1$ variables.

Si la proyección $X_{x_1 \ldots x_{n-1}}$ del conjunto X sobre el hiperplano R^{n-1} puede, a su vez, representarse en una forma análoga a la del conjunto X, entonces la integral $(n-1)$-múltiple, obtenida en el segundo miembro de la igualdad (45.20), puede ser reducida a una integral $(n-2)$-múltiple. Continuando este proceso, si esto es, por supuesto, posible, llegaremos a continuación a una fórmula del tipo

$$\overbrace{\int\cdots\int}^{n \text{ veces}}_{X} f(x_1, \ldots, x_n)dx_1 \ldots dx_n =$$

$$= \int_a^b dx_1 \int_{\varphi_1(x_1)}^{\psi_1(x_1)} dx_2 \int_{\varphi_2(x_1, x_2)}^{\psi_2(x_1, x_2)} dx_3 \ldots$$

$$\ldots \int_{\varphi_{n-1}(x_1, \ldots, x_{n-1})}^{\psi_{n-1}(x_1, \ldots, x_{n-1})} f(x_1, \ldots, x_n)dx_n. \quad (45.21)$$

De este modo, en el caso que se considera la integración de la función de n variables se reduce a la integración sucesiva n veces de las funciones de una variable.

Designemos ahora con $X_{x_1 \ldots x_m}$ la proyección del conjunto X al espacio $R^m_{x_1 \ldots x_m}$, y con $X(x_1, \ldots, x_m)$, la sección del conjunto X mediante los hiperplanos de dimensión $n-m$ que pasan por el punto $(x_1, \ldots, x_m, 0, \ldots, 0)$ y son ortogonales al subespacio $R^m_{x_1 \ldots x_m}$. Al reunir en la fórmula (45.21) las m primeras y $n-m$ últimas integraciones, obtendremos

$$\overbrace{\int\cdots\int}^{n \text{ veces}}_{X} f(x_1, \ldots, x_n)dx_1 \ldots dx_n =$$

$$= \overbrace{\int\cdots\int}^{m \text{ veces}}_{X_{x_1 \ldots x_m}} dx_1 \ldots dx_m \overbrace{\int\cdots\int}^{n-m \text{ veces}}_{X_{(x_1, \ldots, x_m)}} f(x_1, \ldots$$

$$\ldots, x_n)dx_{m+1} \ldots dx_n. \quad (45.22)$$

Si $f(x_1, \ldots, x_n) \equiv 1$ en X, entonces, de esta fórmula obtenemos por analogía con (45.19)

$$\mu X = \int\cdots\int_{X_{x_1 \ldots x_m}} \mu X(x_1, \ldots, x_m)dx_1 \ldots dx_m. \quad (45.23)$$

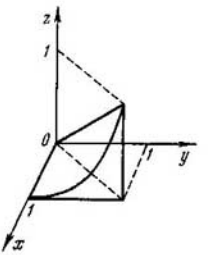


Fig. 192

Ejemplo. Calculemos la integral de la función $f(x, y, z) = xy^2z^3$, extendida a una región finita G, limitada por las superficies $z = xy$, $y = x$, $x = 1$ y $z = 0$ (fig. 192). Al aplicar la fórmula (45.16), tendremos

$$\iiint\limits_G xy^2z^3dxdydz = \int\limits_0^1 xdx \int\limits_0^x y^2dy \int\limits_0^{xy} z^3dz =$$

$$= \frac{1}{4} \int\limits_0^1 x^5dx \int\limits_0^x y^6dy = \frac{1}{28} \int\limits_0^1 x^{12}dx = \frac{1}{364}.$$

Ejercicios. Calcúlense las integrales:

15. $\iiint\limits_X zdxdydz$, $X = \{(x, y, z) : x^2 + y^2 + z^2 \leqslant 1, x^2 + y^2 \leqslant z^2\}$.

16. $\iiint\limits_X (x + y + z)x^2y^2z^2dxdydz$, $X = \{(x, y, z): x \geqslant 0; y \geqslant 0; z \geqslant 0;$ $x + y + z \leqslant 1\}$.

17. $\iiint\limits_X (4x - y + z)dxdydz$; la región X está acotada por las partes de las superficies $x = 0, y = 0, z = 0, x + y = 1, z = 2 - x^2$.

18. $\iiint\limits_X z^2dxdydz$; la región X es la parte común de las bolas $x^2 + y^2 + z^2 \leqslant R^2$ y $x^2 + y^2 + z^2 \leqslant 2Rz$.

45.3*. DESIGUALDAD INTEGRAL GENERALIZADA DE MINKOWSKI

Como otro ejemplo de aplicación de la regla de cambio del orden de integración, demostremos una desigualdad integral de uso frecuente.

Supongamos que la función $f(x, y)$ es continua en el rectángulo $\Delta = \{(x, y) : a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$. Evidentemente, para cualquier $y \in [c, d]$ fijo es continua respecto de x en el segmento $[a, b]$, y para cualquier $x \in [a, b]$ fijo es continua

respecto de y en el segmento $[c, d]$.

Para todo $p > 1$ resulta válida la *desigualdad generalizada de Minkowski*

$$\left\{\int_c^d \left[\int_a^b |f(x, y)|\,dx\right]^p dy\right\}^{1/p} \leqslant \int_a^b dx \left[\int_c^d |f(x, y)|^p\,dy\right]^{1/p} \tag{45.24}$$

Pongamos

$$F(y) = \int_a^b |f(x, y)|\,dx. \tag{45.25}$$

La función F es continua (véase el lema 1 en el p. 45.1) y no negativa en el segmento $[c, d]$. Por ello, su p-ésima potencia es también integrable y no negativa en este segmento, y $0 \leqslant \int_c^d F^p(y)dy < +\infty$.

Si $\int_c^d F^p(y)dy = 0$, entonces, en virtud de la continuidad de la función F, tendremos (véase la propiedad 9 en el p. 28.1): $F(y) \equiv 0$ en $[c, d]$. Por eso de la fórmula (45.25), debido a la misma propiedad, se deduce que para cualquier $y \in [c, d]$ tiene lugar $f(x, y) \equiv 0$ en $[a, b]$, es decir, $f(x, y) \equiv 0$ en Δ. En este caso la desigualdad (45.24) es, evidentemente, válida.

Sea $\int_c^d F^p(y)dy > 0$. Al cambiar el orden de integración y al aplicar la desigualdad de Hölder (28.48), obtendremos, en virtud de (45.25),

$$\int_c^d F^p(y)dy = \int_c^d F^{p-1}(y)\left[\int_a^b |f(x, y)|\,dx\right]dy =$$

$$= \int_a^b dx \int_c^d |f(x, y)|F^{p-1}(y)dy \leqslant$$

$$\leqslant \int_a^b \left[\int_c^d |f(x, y)|^p dy\right]^{1/p}\left[\int_c^d F^{q(p-1)}(y)dy\right]^{1/q} dx, \tag{45.26}$$

donde $\frac{1}{p} + \frac{1}{q} = 1$, y, por lo tanto, $q(p-1) = p$. Al reducir ambos miembros de la igualdad (45.26) por el factor $\left(\int_c^d F^p(y)\,dy\right)^{1/q} \neq 0$, tendremos

$$\left(\int_c^d F^p(y)dy\right)^{1/p} \leqslant \int_a^b \left[\int_c^d |f(x, y)|^p dy\right]^{1/p} dx.$$

Sustituyendo aquí (45.25), obtenemos la desigualdad (45.24). □

La condición de continuidad de la función f no es esencial para que la desigualdad (45.24) sea válida y puede ser debilitada. Con el fin de simplificar la demostración, a título de dominio de definición de la función f se ha empleado un rectángulo. Para las suposiciones más generales, la demostración de la desigualdad de Minkowski, basada sobre la misma idea, puede encontrarse en la monografía "Desigualdades" por Hardy G. H.; Littlewood D. E., Polya G. (Hardy G. H., Littlewood D. E., Polya, Inequalities, 1934).

§ 46. CAMBIO DE VARIABLES EN UNA INTEGRAL MÚLTIPLE

46.1. SENTIDO GEOMÉTRICO DEL MÓDULO DE JACOBIANO EN EL CASO BIDIMENSIONAL

Supongamos que G es un conjunto abierto en el plano R^2_{uv} y G^*, un conjunto abierto en el plano R^2_{xy}, F es la aplicación de G sobre G^* y

$$M = (u, v) \in G, \quad M^* = (x, y) \in G^*, \quad F(M) = M^*.$$

La aplicación F se da mediante un par de funciones

$$x = x(u, v), \quad y = y(u, v). \tag{46.1}$$

Supondremos que F satisface las siguientes condiciones:

1) aplica biunívocamente G sobre G^*;

2) es continuamente derivable en G;

3) el jacobiano $J(u, v) = \dfrac{\partial(x, y)}{\partial(u, v)}$ no se reduce a cero en G.

Observemos que la aplicación F^{-1}, inversa de F, es también una aplicación biunívoca continuamente derivable provista de un jacobiano distinto de cero en G^* (véase el p. 41.7). Es por eso, en particular, que F es una aplicación difeomorfa del conjunto abierto G (véase la definición 11 en el p. 41.7) sobre G^*.

Si γ es un contorno sencillo cerrado dispuesto en G, entonces, en vista de que la aplicación F es biunívoca, su imagen $\gamma^* = F(\gamma)$ es también un contorno sencillo cerrado.

Lema 1. *Sea Γ un conjunto acotado abierto y $\bar{\Gamma} \subset G$. En este caso $\Gamma^* = F(\Gamma)$ es también un conjunto acotado abierto y*

$$\partial F(\Gamma) = F(\partial\Gamma). \tag{46.2}$$

DEMOSTRACIÓN. Como F y F^{-1} son aplicaciones homeomorfas, los conjuntos abiertos en estas mismas se aplican en conjuntos abiertos. Por consiguiente, los puntos interiores de un conjunto cualquiera, Γ, por ejemplo, o Γ^*, respectivamente, pasan a los puntos interiores de su imagen y los puntos de frontera, a los de frontera.

En efecto, sea, por ejemplo, M un punto interior del conjunto Γ, es decir, existe un entorno del punto $U = U(M)$ dispuesto en Γ: $U \subset \Gamma$. Entonces, el entorno

$U^* = F(U)$ del punto $M^* = F(M)$ se dispone en Γ^*: $U^* \subset \Gamma^*$, es decir, M^* es un punto interior del conjunto Γ^*[1]).

Sea ahora M un punto de frontera del conjunto Γ, $M^* = F(M)$ y U^* es un entorno del punto M^*. Por ser homeomorfa la aplicación F, el conjunto $U = F^{-1}(F^*)$ es el entorno del punto M, y como $M \in \partial\Gamma$, entonces en el entorno U se tienen tanto los puntos que pertenecen al conjunto Γ, como los que no le pertenecen. Por consiguiente, en el entorno U^* del punto $M^* = F(M)$ (teniendo en cuenta que este entorno es la imagen del entorno $U = U(M)$ del punto M cuando se realiza la aplicación F) hay también tanto puntos pertenecientes al conjunto Γ^* como aquellos que no le pertenecen, es decir, los puntos de frontera realmente se aplican en puntos de frontera:

$$F(\partial\Gamma) \subset \partial\Gamma^*. \tag{46.3}$$

Ya que los razonamientos análogos son lícitos también para la aplicación inversa, en la fórmula (46.3) podemos sustituir el signo de inclusión por el de igualdad, es decir, se cumple la condición (46.2). Además, de lo que el conjunto Γ es abierto se deduce, en virtud de lo demostrado, el carácter abierto del conjunto Γ^*. Luego, dado que Γ es un conjunto acotado, lo será también el conjunto cerrado $\bar{\Gamma}$. Por ello, de acuerdo con el lema 3 del p. 41.4., el conjunto $F(\bar{\Gamma})$ está acotado. Del carácter acotado del conjunto $F(\bar{\Gamma})$ se desprende que el conjunto $\Gamma^* = F(\Gamma)$ es también acotado, pues $F(\Gamma) \subset F(\bar{\Gamma})$. □

Corolario. *Si la frontera $\partial\Gamma$ en las suposiciones del lema 1 se compone de un número finito de las curvas derivables continuamente a trozos, los conjuntos abiertos Γ y Γ^* son cuadrables.*

DEMOSTRACIÓN. Si γ es una curva continuamente derivable dispuesta en G y $u = u(t)$, $v = v(t)$, $a \leqslant t \leqslant b$, es una representación de dicha curva, las funciones $u(t)$ y $v(t)$ son continuamente derivables en el segmento $[a, b]$. En la aplicación F la curva γ pasará en la curva $\gamma^* = F(\gamma)$ cuya representación será

$$x(t) = x(u(t), v(t)), \quad y(t) = y(u(t), v(t)), \quad a \leqslant t \leqslant b,$$

donde, en virtud de las fórmulas de derivación de una función compuesta (véase el p. 20.3) y del teorema sobre la continuidad de la composición de funciones continuas (véase el p. 19.5), las funciones $x(t)$ e $y(t)$ tienen también derivadas continuas en el segmento $[a, b]$. Por consiguiente, la curva γ^* es también continuamente derivable. De aquí se infiere inmediatamente que si γ es una curva derivable continuamente a trozos, es decir, es una reunión de un número finito de curvas continuamente derivables (véase el p. 16.3), entonces γ^* es también una curva derivable continuamente a trozos.

Ahora, si la frontera $\partial\Gamma$ de un conjunto abierto $\Gamma \subset G$ se compone de un número finito de las curvas derivables continuamente a trozos, la frontera $\partial\Gamma^*$ del conjunto abierto $\Gamma^* \subset G^*$ consta también, en virtud de lo dicho anteriormente, de un

---

[1]) Hemos obtenido esta afirmación sólo como una consecuencia directa del homeomorfismo de la aplicación F. Desde luego, en el caso dado esto se deduce directamente de las suposiciones más fuertes admitidas más arriba (véase el corolario del teorema 7 en el p. 41.8).

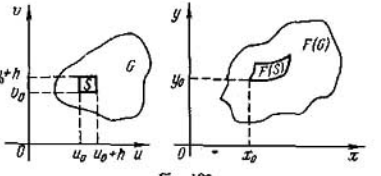


Fig. 193

número finito de curvas derivables continuamente a trozos. Por consiguiente, tanto $\partial\Gamma$, como $\partial\Gamma^*$ son rectificables (véase el teorema 1 en el p. 16.5), a consecuencia de lo cual tienen medida cero (véase el teorema 4 en el p. 44.2). Por ello, en el caso considerado los conjuntos abiertos Γ y Γ^* son cuadrables, pues tienen las fronteras de medida cero. □

Supongamos ahora que $(u_0, v_0) \in G$ y h es cierto número. Examinemos un cuadrado cerrado S (fig. 193) cuyos vértices están dispuestos en los puntos

$$(u_0, v_0), (u_0 + h, v_0), (u_0 + h, v_0 + h), (u_0, v_0 + h). \tag{46.4}$$

Sea $S \subset G$ (para h suficientemente pequeño esta inclusión siempre se cumple; ¿por qué?). La frontera ∂S del cuadrado S compuesta de sus cuatro lados es, obviamente, un contorno cerrado sencillo suave a trozos. En virtud del corolario del lema 2, el conjunto $S^* = F(S)$ (véase fig. 193) representa una *región* cuadrable cerrada (que S^* es una región cerrada se desprende del principio de conservación de la región, véase el p. 41.8).

Veamos cómo se porta la razón

$$\mu F(S)/\mu S^* \tag{46.5}$$

cuando h tiende a cero.

Introduzcamos las designaciones:

$$x(u_0, v_0) = x_0, \quad y(u_0, v_0) = y_0,$$

$$\frac{\partial x(u_0, v_0)}{\partial u} = a_{11}, \quad \frac{\partial x(u_0, v_0)}{\partial v} = a_{12}, \quad \frac{\partial y(u_0, v_0)}{\partial u} = a_{21}, \quad \frac{\partial y(u_0, v_0)}{\partial v} = a_{22},$$

$$u - u_0 = \Delta u, \quad v - v_0 = \Delta v, \quad r = \sqrt{\Delta u^2 + \Delta v^2}.$$

Por ser derivables las funciones (46.1), son válidas las siguientes fórmulas

$$\begin{aligned} x &= x(u, v) = x_0 + a_{11}(u - u_0) + a_{12}(v - v_0) + \varepsilon_1 r, \\ y &= y(u, v) = y_0 + a_{21}(u - u_0) + a_{22}(v - v_0)\varepsilon_2 r, \end{aligned} \tag{46.6}$$

donde las funciones $\varepsilon_i = \varepsilon_i(u_0, v_0, \Delta u, \Delta v)$, $i = 1, 2$ tienden a cero cuando $r \to 0$.

*) Aquí, como siempre, μX denota la medida (el área, en el caso dado) del conjunto X.

A la par con la aplicación F, consideraremos una aplicación lineal $\tilde{F}$ del plano R^2_{uv} sobre el plano R^2_{xy}, definida por las fórmulas

$$\begin{aligned} \tilde{x} &= x_0 + a_{11}(u - u_0) + a_{12}(v - v_0), \\ \tilde{y} &= y_0 + a_{21}(u - u_0) + a_{22}(v - v_0), \end{aligned} \tag{46.7}$$

Por el curso de geometría analítica se sabe que, realizándose una aplicación lineal, la imagen de todo paralelogramo y, en particular, de un cuadrado es un paralelogramo, con la particularidad de que la razón entre el área del último y el área del paralelogramo que se aplica es igual al valor absoluto del determinante de la aplicación el cual, en el caso de la aplicación $\tilde{F}$, coincide con el jacobiano $J(u, v)$ de la aplicación F en el punto (u_0, v_0). De este modo, en el caso dado, para la aplicación (46.7) tenemos

$$\frac{\mu\tilde{F}(S)}{\mu S} = \left| \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right| = |J(u_0, v_0)|. \tag{46.8}$$

La aplicación continuamente derivable F en el entorno del punto (u_0, v_0) difiere de la aplicación lineal $\tilde{F}$ en una función infinitamente pequeña de orden superior que el incremento de los argumentos (véase (46.6)). Probemos que de aquí se infiere la validez de la igualdad

$$\lim_{h \to 0} \frac{\mu F(S)}{\mu S} = |J(u_0, v_0)|. \tag{46.9}$$

Más aún, mostremos que la razón en esta igualdad tiende a su límite uniformemente en cualquier compacto dispuesto en el conjunto abierto G. Enunciemos este resultado como un teorema.

Teorema 1. *Supongamos que la aplicación F de un conjunto abierto $G \subset R^2_{uv}$ sobre otro conjunto abierto $G^* \subset R^2_{xy}$ es biunívoca y continuamente derivable en G y supongamos también que el jacobiano $J(u, v)$ no se reduce a cero en G. Entonces, si S es un cuadrado cuyos vértices coinciden con los puntos* (46.4), *se tiene*

$$\frac{\mu F(S)}{\mu S} = |J(u_0, v_0)| + \varepsilon(u_0, v_0, h), \tag{46.10}$$

*donde la función $\varepsilon = \varepsilon(u_0, v_0, h)$, para $h \to 0$, tiende a cero uniformemente respecto a (u_0, v_0) en cualquier compacto $A \subset G$.**)

Corolario. *Para todo punto (u_0, v_0) de un conjunto abierto G se verifica la igualdad* (46.9).

DEMOSTRACIÓN. Probemos que el área de la imagen del cuadrado S, realizándose la aplicación F, se diferencia del área de la imagen de este cuadrado, cuando se realiza la aplicación $\tilde{F}$, en un infinitésimo de orden más elevado que el área h^2 del mismo cuadrado S y esta estimación es uniforme en cualquier compacto $A \subset G$, es decir, que

$$\mu F(S) = \mu\tilde{F}(S) + \varepsilon h^2, \tag{46.11}$$

---

*) De este modo, $A \ni (u_0, v_0)$.

donde ε tiende a cero uniformemente en el conjunto A, cuando la longitud h del lado del cuadrado S tiende a cero (véase en el p. 39.4. la definición que muestra cómo una función tiende al límite). Por cuanto (véase (46.8))

$$\mu\tilde{F}(S) = |J(u_0, v_0)|\mu S \tag{46.12}$$

y

$$\mu(S) = h^2,$$

de (46.11) se desprende directamente la afirmación del teorema, es decir, la fórmula (46.10).

Pasando a la demostración de la fórmula (46.11), fijemos, ante todo, el conjunto A. Ya que A es un compacto y $A \subset G$, las funciones $\varepsilon_i = \varepsilon_i(u_0, v_0, \Delta u, \Delta v)$, $i = 1, 2$, (véase (46.6)) tienden uniformemente hacia cero en el conjunto A para $r \to 0$ (véase la observación al teorema 4 en el p. 20.2, como también el p. 39.4). Los conjuntos A y $X^2_{uv} \setminus G$ no se intersecan y son cerrados, siendo A, además, acotado, a consecuencia de lo cual (véase el lema 7 en el p. 18.2) $\eta = \rho(A, X^2_{uv} \setminus G) > 0$.

En adelante escogeremos h de modo tal que sea $|h| < \dfrac{\eta}{\sqrt{2}}$. En este caso de lo que $(u_0, v_0) \in A$ proviene que $S \subset G$.

Estimemos la distancia entre las imágenes de un mismo punto del cuadrado S en las aplicaciones F y $\tilde{F}$. Sea

$$M = (u, v) \in S,\ F(M) = (x, y) \quad \text{y} \quad \tilde{F}(M) = (\xi, \eta).$$

De (46.6) y (46.7) obtenemos $x = \xi + \varepsilon_1 r$, $y = \eta + \varepsilon_2 r$, y por ende,

$$\rho(F(M), \tilde{F}(M)) = \sqrt{(x - \xi)^2 + (y - \eta)^2} = r\sqrt{\varepsilon_1^2 + \varepsilon_2^2}.$$

Como r es la distancia entre el vértice (u_0, v_0) del cuadrado S y el punto $M \in S$, y $|h|\sqrt{2}$ es la longitud de la diagonal del cuadrado S, entonces, evidentemente, se verifica la desigualdad $r \leqslant |h|\sqrt{2}$ y por esta razón se tiene

$$d = \sup_{M \in S} \rho(F(M), \tilde{F}(M)) \leqslant |h|\varepsilon_3(u_0, v_0, h), \tag{46.13}$$

donde $\varepsilon_3 = \varepsilon_3(u_0, v_0, h) = \sup_{M \in S} \sqrt{2(\varepsilon_1^2 + \varepsilon_2^2)}$ tiende a cero uniformemente en el conjunto A, para $h \to 0$.

Construyamos los paralelogramos cerrado $\bar{S}_e$*) y abierto $\bar{S}_i$**) cuyos lados sean paralelos a los del paralelogramo $\bar{S} = \tilde{F}(S)$ y disten a la magnitud d de los lados correspondientes suyos (fig. 194) de modo tal que sea

$$\bar{S}_i \subset \bar{S} = \tilde{F}(S) \subset \bar{S}_e. \tag{46.14}$$

Probemos ante todo que, siendo h suficientemente pequeños, el conjunto $\bar{S}_i$ no es vacío. Más aún, mostremos que el paralelogramo $\bar{S}_i$ contiene en sí un círculo abierto de radio d cuyo centro se ubica en el centro del paralelogramo S.

*) "e" es la letra inicial de la palabra exterior.

**) "i" es la letra inicial de la palabra interior.

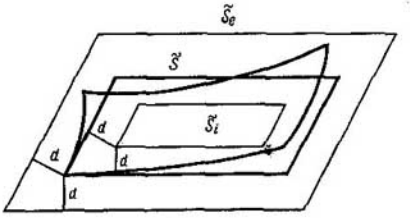


Fig. 194

Designemos mediante a y b las longitudes de los lados del paralelogramo $\tilde{S}$ y mediante H_a y H_b, las longitudes de sus alturas tiradas en los lados de longitud a y b, respectivamente (fig. 195). Para demostrar que, siendo h suficientemente pequeños, el círculo de radio d y centro ubicado en el centro del paralelogramo $\tilde{S}$ está contenido en $\tilde{S}_i$, basta, obviamente, establecer la validez, para h suficientemente pequeños, de las desigualdades

$$4d < H_a \quad 4d < H_b. \tag{46.15}$$

Demostrémoslo. Supongamos, para concretar, que el lado del paralelogramo $\tilde{S}$ de longitud a une los vértices que en la aplicación F constituyen las imágenes de los vértices (u_0, v_0) y $(u_0 + h, v_0)$ del cuadrado S, es decir, une los puntos (x_0, y_0) y $(x_0 + a_{11}h, y_0 + a_{21}h)$. En este caso

$$a = \sqrt{a_{11}^2 h^2 + a_{21}^2 h^2} = |h| \sqrt{a_{11}^2 + a_{21}^2}. \tag{46.16}$$

Análogamente,

$$b = |h| \sqrt{a_{12}^2 + a_{22}^2}. \tag{46.17}$$

Las funciones $a_{ij} = a_{ij}(u_0, v_0)$, $i, j = 1, 2$, son valores de las derivadas parciales correspondientes de las funciones $x(u, v)$ e $y(u, v)$ en los puntos (u_0, v_0) del compacto A. En virtud de la continuidad supuesta de estas derivadas parciales, ellas

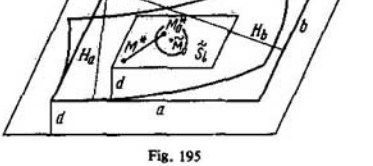


Fig. 195

están acotadas en el conjunto A, es decir, existe tal constante $c_1 > 0$, que en A se verifican las desigualdades $|a_{ij}| \leqslant c_1$, $i, j = 1, 2$.

De aquí y de las fórmulas (46.16) y (46.17) se deduce que

$$|a| \leqslant c_1\sqrt{2}\,|h|, \tag{46.18}$$

$$|b| \leqslant c_1\sqrt{2}\,|h|. \tag{46.19}$$

Luego, por hipótesis, el jacobiano $J(u, v)$ de la aplicación F, que representa una función continua, no se reduce a cero en el conjunto G y, por consiguiente, tampoco en el compacto A. Por esta razón existe (¿por qué?) una constante $c_2 > 0$ tal que en el conjunto A se verifica la desigualdad

$$|J(u, v)| \geqslant c_2. \tag{46.20}$$

Al notar que $\mu S = aH_a = bH_b = |J(u_0, v_0)|\,h^2$, obtendremos (véanse (46.18), (46.19) y (46.20)):

$$h^2 = \frac{aH_a}{|J(u_0, v_0)|} \leqslant \frac{1}{c_2}\, c_1\sqrt{2}\,|h|\,H_a,$$

$$h^2 = \frac{bH_b}{|J(u_0, v_0)|} \leqslant \frac{1}{c_2}\, c_1\sqrt{2}\,|h|\,H_b,$$

es decir,

$$|h| \leqslant \frac{c_1\sqrt{2}}{c_2} H_a, \tag{46.21}$$

$$|h| \leqslant \frac{c_1\sqrt{2}}{c_2} H_b. \tag{46.22}$$

Disponiendo de estas estimaciones, es fácil demostrar la validez de las desigualdades (46.15). En efecto, al hacer uso de las desigualdades (46.13), (46.21) y (46.22), obtendremos

$$4d \leqslant 4\varepsilon_3|h| \leqslant \frac{4\sqrt{2}c_1\varepsilon_3}{c_2} H_a, \tag{46.23}$$

$$4d \leqslant \frac{4\sqrt{2}c_1\varepsilon_3}{c_2} H_b. \tag{46.24}$$

Elijamos ahora tal $\delta > 0$ que para $|h| < \delta$ y $(u_0, v_0) \in A$ se cumpla la condición

$$\frac{4\sqrt{2}c_1\varepsilon_3}{c_2} < 1. \tag{46.25}$$

Esto es siempre factible debido a que la función $\varepsilon_3 = \varepsilon_3(u_0, v_0, h)$ (véase (46.13)) tiende a cero uniformemente en el compacto A cuando $h \to 0$. De (46.23), (46.24) y (46.25) se desprende que para $|h| < \delta$ se verifican las desigualdades (46.15), de donde, en particular, se deduce que el conjunto S_i no es vacío. En lo que sigue supondremos siempre, durante la demostración, que $|h| < \delta$.

El conjunto $\bar{S}_e \setminus \bar{S}_i$ se denominará *cuadro* y se denotará con $\bar{R}$:

$$\bar{R} = \bar{S}_e \setminus \bar{S}_i.$$

El cuadro $\bar{R}$ es un conjunto cerrado, pues $\bar{S}_e$ es un paralelogramo cerrado y $\bar{S}_i$, abierto. Representa una reunión de cuatro trapecios, privados de puntos interiores comunes, cuyas alturas miden $2d$ y las líneas medias coinciden con los lados correspondientes del paralelogramo S. Por ello, $\mu\bar{R} = 4d(a + b)$.

Observemos que si el conjunto $\bar{S}_i$ fuera vacío, el cálculo del área del cuadro $\bar{R}$ se realizaría de otro modo: los trapecios mencionados se transformarían en unos triángulos en los cuales los lados del paralelogramo ya no serían, por regla general, las líneas medias.

De la expresión obtenida para el área $\mu\bar{R}$ del cuadro $\bar{R}$ se deduce, en virtud de las desigualdades (46.13), (46.18) y (46.19), que $\mu\bar{R} \leqslant 8\sqrt{2}c_1\varepsilon_3 h^2$. Al poner $\varepsilon_4 = 8\sqrt{2}c_1\varepsilon_3$, tendremos en definitiva

$$\mu\bar{R} \leqslant \varepsilon_4 h^2, \tag{46.26}$$

donde la función ε_4 tiende uniformemente hacia cero en el compacto A cuando $h \to 0$.

Mostremos ahora que el área del conjunto $F(S)$ se diferencia del área del paralelogramo $\bar{S} = \bar{F}(S)$ en una magnitud no superior al área del cuadro $\bar{R}$. Para ello establezcamos ante todo que

$$\bar{S}_i \subset F(S) \subset \bar{S}_e. \tag{46.27}$$

En efecto, si $M \in S$, entonces $\bar{F}(M) \in \bar{S}$ y, de conformidad con (46.13), $\rho(F(M), \bar{F}(M)) \leqslant d$.

Luego, el conjunto $\bar{S}_e$ contiene, por construcción, todos los puntos del plano que distan del paralelogramo $\bar{S}$ a una magnitud no superior a d. Por esta razón $F(M) \in \bar{S}_e$ y la inclusión $F(S) \subset \bar{S}_e$ queda demostrada. Resta demostrar que $\bar{S}_i \subset F(S)$. Hemos de notar, ante todo, que

$$F(\partial S) \subset \bar{R}. \tag{46.28}$$

En efecto, si $M \in \partial S$, se tiene $F(M) \in \partial\bar{S}$ y, de acuerdo con (46.13), $\rho(F(M), \bar{F}(M)) \leqslant d$. Pero el cuadro $\bar{R}$ contiene, por construcción, todos los puntos del plano que distan de la frontera $\partial\bar{S}$ del paralelogramo $\bar{S}$ a una magnitud no superior a d, por lo cual $F(M) \in \bar{R}$ y la inclusión (46.28) queda demostrada. Como, bajo las suposiciones asumidas, la frontera $\partial F(S)$ de la imagen $F(S)$ del cuadrado S coincide con la imagen $F(\partial S)$ de la frontera ∂S del cuadrado S (véase el lema 1), la inclusión (46.28) puede escribirse en la forma

$$\partial F(S) \subset \bar{R}. \tag{46.29}$$

Sea ahora M_0 el centro del cuadrado S. En la aplicación $\bar{F}$ éste pasa al centro $\bar{M}_0 = \bar{F}(M_0)$ del paralelogramo $\bar{S}$. Sea Q un círculo cerrado de radio d con centro en el punto $\bar{M}_0$ (la magnitud d se determina por la fórmula (46.13)). Hemos demostrado anteriormente que $Q \subset \bar{S}_i$. Si $M_0^* = F(M_0)$, entonces, de acuerdo con (46.13), $\rho(M_0^*, \bar{M}_0) \leqslant d$, y, por lo tanto, $M_0^* \in Q$, por lo cual también $M_0^* \in \bar{S}_i$. De este modo, la clausura $\bar{S}_i$ contiene, a ciencia cierta, un punto de $F(S)$, a saber, la imagen M_0^* del centro M_0 del cuadrado S en la aplicación F.

Probemos ahora que todos los puntos $\bar{S}_i$ también pertenecen a $F(S)$. Admitamos lo contrario; supongamos que existe un punto $M^* \in \bar{S}_i$ tal que $M^* \notin F(S)$ (véase la fig. 195). Todo segmento es, obviamente, un conjunto linealmente conexo, por lo cual, de acuerdo con el lema 9 del p. 18.2, en el segmento $M_0^* M^*$ con sus extremos en los puntos M_0^* y M^* hay un punto de la frontera $\partial F(S)$ del conjunto $F(S)$. De este punto no puede servir M_0^*, puesto que M_0 es un punto interior del conjunto S y para la aplicación F los puntos interiores pasan a interiores, por eso el punto $M_0^* = F(M_0)$ es interior y no de frontera del conjunto $F(S)$. Debido a esto el punto de intersección del conjunto $\partial F(S)$ con el segmento $M_0^* M^*$ es un punto del paralelogramo abierto $\bar{S}_i$, lo que contradice la inclusión (46.29).

De este modo, no existe un punto $M^* \in \bar{S}_i$, para el cual sea, a la vez, $M^* \notin F(S)$, por lo cual $\bar{S}_i \subset F(S)$. La fórmula (46.27) se ha demostrado.

De (46.14) y (46.27) proviene que

$$\mu \bar{S}_i \leqslant \mu \bar{F}(S) \leqslant \mu \bar{S}_e, \quad \mu \bar{S}_i \leqslant \mu F(S) \leqslant \mu \bar{S}_e,$$

y, por lo tanto,

$$|\mu F(S) - \mu \bar{F}(S)| \leqslant \mu \bar{S}_e - \mu \bar{S}_i = \mu \bar{R}.$$

Por esta razón, en vista de (46.26)

$$|\mu F(S) - \mu \bar{F}(S)| \leqslant \varepsilon_4 h^2, \tag{46.30}$$

donde ε_4 tiende hacia cero uniformemente en el compacto A cuando $h \to 0$.

Pongamos

$$\varepsilon(u_0, v_0, h) = \frac{\mu F(S) - \mu \bar{F}(S)}{h^2}, \tag{46.31}$$

entonces de (46.30) se infiere que $|\varepsilon| \leqslant |\varepsilon_4|$ y, por ende, en el conjunto A ε tiende a cero uniformemente cuando $h \to 0$. De (46.31) tenemos

$$\mu F(S) = \mu \bar{F}(S) + \varepsilon h^2,$$

es decir, se ha obtenido, la fórmula (46.11), de donde, como ya se ha señalado, se deduce directamente (46.10). □

46.2. CAMBIO DE VARIABLES EN UNA INTEGRAL DOBLE

Dejamos intactas al principio las designaciones y suposiciones asumidas en el punto antecedente, en particular, supondremos que F es una aplicación biunívoca continuamente derivable del conjunto abierto $G \subset R^2_{uv}$ sobre otro conjunto abierto $G^* \subset R^2_{xy}$, provisto de un jacobiano distinto de cero en G. Sean Γ y Γ^* unos conjuntos abiertos cuadrables (y por eso acotados), $\bar{\Gamma} \subset G$, $\bar{\Gamma}^* \subset G^*$, y supongamos que, al realizarse la aplicación F, el conjunto $\bar{\Gamma}$ se aplica sobre $\bar{\Gamma}^*$. En este caso $\bar{\Gamma}$ y $\bar{\Gamma}^*$ son unos compactos, los puntos interiores de Γ pasan a puntos interiores, mientras que la frontera de Γ se aplica sobre la de Γ^*.

Teorema 2 (fórmula de cambio de las variables en una integral doble). *Sea $f(x, y)$ una función definida y continua en $\bar{\Gamma}^*$. En este caso*

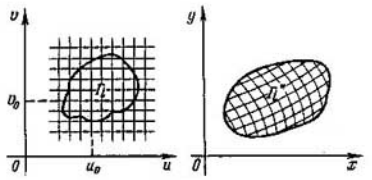


Fig. 196

$$\iint\limits_{\Gamma^*} f(x, y)\, dxdy = \iint\limits_{\Gamma} f[(x(u, v), y(u, v)]\left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv. \quad (46.32)$$

DEMOSTRACIÓN. Observemos que las integrales que figuran en (46.32) existen como integrales de las funciones continuas en la clausura de las regiones cuadrables. Efectivamente, por hipótesis, la función $f(x, y)$ es continua en $\overline{\Gamma^*}$ y el jacobiano $\dfrac{\partial(x, y)}{\partial(u, v)}$, en $\overline{\Gamma}$, mientras que la función $f[x(u, v), y(u, v)]$ es continua en $\overline{\Gamma}$ como composición de las funciones continuas.

Tomemos una partición de rango k del plano R_{uv} en cuadrados. Elijamos el rango k tan grande que todo cuadrado de este rango que se interseca con $\overline{\Gamma}$ se contenga íntegramente en G (¿por qué tal rango existe?). Designemos con Γ_i, $i = 1, 2, \ldots, i_k$ toda clase de intersecciones no vacías de los interiores (los conjuntos de puntos interiores) de los cuadrados de rango k con el conjunto Γ. Los conjuntos Γ_i son cuadrables y abiertos, pues sus fronteras tienen medida cero, puesto que se componen, en el caso general, de una parte de la frontera del cuadrado correspondiente de rango k y una parte de la frontera del conjunto Γ. La totalidad $\tau_k = \{\Gamma_i\}_{i=1}^{i=i_k}$ forma la partición del conjunto Γ, con la particularidad de que, evidentemente,

$$\lim_{k\to\infty} \delta_{\tau_k} = 0. \quad (46.33)$$

Sea ahora $\Gamma_i^* = F(\Gamma_i)$; la frontera de Γ_i se aplica sobre la frontera de Γ_i^*, razón por la cual la frontera de Γ_i^* se compone, en el caso general, de una parte de la frontera del conjunto Γ^* (al admitir que el conjunto Γ^* es cuadrable, convenimos en que esta frontera tiene medida cero) y una parte de una curva suave a trozos que es la imagen de la frontera del cuadrado correspondiente y tiene también, debido a esta circunstancia, medida cero. De lo dicho proviene que Γ_i^* es un conjunto abierto cuadrable. De lo que la aplicación F es biunívoca se deduce que la totalidad $\tau_k^* = \{\Gamma_i^*\}_{i=1}^{i=i_k}$ forma la partición del conjunto Γ^* (fig. 196).

Estimemos la finura de la partición τ_k^*. Sea δ_k el diámetro del cuadrado de rango k $\left(\text{evidentemente, } \delta_k = \dfrac{\sqrt{2}}{10^k}\right)$ y $M_1^* = (x_1, y_1) \in \Gamma_i^*$, $M_2^* = (x_2, y_2) \in \Gamma_i^*$. En este caso

existen tales $M_1 \in \Gamma_i$ y $M_2 \in \Gamma_j$ que $F(M_1) = M_1^*, F(M_2) = M_2^*$, con la particularidad de que $\rho(M_1, M_2) < \delta_k$. Por consiguiente,

$$\rho(M_1^*, M_2^*) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \leqslant \sqrt{\omega^2(\delta_k; x) + \omega^2(\delta_k; y)}, \quad (46.34)$$

donde $\omega(\delta; x)$ y $\omega(\delta; y)$ son los módulos de continuidad de las funciones $x = x(u, v)$ e $y = y(u, v)$ en el compacto $\overline{\Gamma}$. Siendo estas funciones continuas en $\overline{\Gamma}$, son uniformemente continuas, por lo cual (véase el p. 19.7)

$$\lim_{k \to \infty} \omega(\delta_k; x) = \lim_{k \to \infty} \omega(\delta_k; y) = 0. \quad (46.35)$$

De (46.34) para el diámetro $d(\Gamma_i^*)$ obtenemos

$$d(\Gamma_i^*) = \sup_{\substack{M_1^* \in \Gamma_i^* \\ M_2^* \in \Gamma_i^*}} \rho(M_1^*, M_2^*) \leqslant \sqrt{\omega^2(\delta_k; x) + \omega^2(\delta_k; y)}.$$

y, por consiguiente,

$$\delta_{\tau_k^*} = \sup_{\Gamma_i^* \in \tau_k^*} d(\Gamma_i^*) \leqslant \sqrt{\omega^2(\delta_k; x) + \omega^2(\delta_k; y)},$$

por lo cual, en virtud de (46.35),

$$\lim_{k \to \infty} \delta_{\tau_k^*} = 0^*). \quad (46.36)$$

Tomemos ahora sólo aquellos elementos de las particiones τ_k y τ_k^*, cuyas clausuras no se intersecan con las fronteras $\partial\Gamma$ y $\partial\Gamma^*$ de los conjuntos Γ y Γ^*. Designémoslos mediante $\tau_k(\partial\Gamma)$ y $\tau_k^*(\partial\Gamma^*)$, respectivamente:

$$\tau_k(\partial\Gamma) = \{\Gamma_i : \Gamma_i \in \tau_k, \overline{\Gamma}_i \cap \partial\Gamma = \varnothing\},$$
$$\tau_k^*(\partial\Gamma^*) = \{\Gamma_i^* : \Gamma_i^* \in \tau_k^*, \overline{\Gamma}_i^* \cap \partial\Gamma^* = \varnothing\}.$$

En virtud de las suposiciones asumidas, $\Gamma_i^* \in \tau_k^*(\partial\Gamma^*)$ cuando, y sólo cuando, $\Gamma_i \in \tau_k(\partial\Gamma)$; en este caso $\tau_k(\partial\Gamma)$ consta de aquellos, y sólo aquellos elementos $\Gamma_i \in \tau_k$ que representan los cuadrados íntegros contenidos junto con sus fronteras en el conjunto Γ.

Formemos las sumas integrales incompletas $\sigma_{\tau_k^*(\partial\Gamma^*)}$ (véase el p. 44.3) para la función $f(x, y)$, tomando a título de puntos $(\xi_i, \eta_i) \in \overline{\Gamma}_i \in \tau_k^*(\partial\Gamma^*)$ las imágenes de los vértices cualesquiera (u_i, v_i) de los cuadrados correspondientes Γ_i:

$$\xi_i = x(u_i, v_i) \quad \eta_i = y(u_i, v_i). \quad (46.37)$$

En otras palabras, consideraremos las sumas de la forma

$$\sigma_{\tau_k^*(\partial\Gamma^*)} = \sum_{\Gamma_i^* \in \tau_k^*(\partial\Gamma^*)} f(\xi_i, \eta_i) \mu\Gamma_i^*. \quad (46.38)$$

---

*) No es difícil convencerse de que la igualdad (46.36) puede obtenerse inmediatamente de la continuidad uniforme de la aplicación continua del compacto (véase el lema 4 en el p. 41.4).

Como se sabe (véase el teorema 5 en el p. 44.3), debido a que la condición (46.36) se cumple, tenemos

$$\lim_{k\to\infty} \sigma_{\tau_k^*(\partial\Gamma^*)} = \iint_{\overline{\Gamma}^*} f(x, y)dxdy. \tag{46.39}$$

Por otra parte, para $\Gamma_i^* = F(\Gamma_i)$, para los cuales Γ_i es un cuadrado, y, por lo tanto, para $\Gamma_{i\in\tau_k}(\partial\Gamma)$, de acuerdo con el teorema 1 del punto anterior, tenemos

$$\mu\Gamma_i^* = |J(u_i, v_i)|\mu\Gamma_i + \varepsilon\mu\Gamma_i, \tag{46.40}$$

donde $\varepsilon = \varepsilon(u_i, v_i, \delta_{\tau_k})$ tiende uniformemente en el compacto $\overline{\Gamma}$ hacia cero cuando $k \to 0$. Sustituyendo (46.37) y (46.40) en (46.38), obtendremos

$$\sigma_{\tau_k^*(\partial\Gamma^*)} = \sum_{\Gamma_i\in\tau_k(\partial\Gamma)} f[x(u_i, v_i), y(u_i, v_i)]\,|J(u_i, v_i)|\mu\Gamma_i + \sum_{\Gamma_i\in\tau_k(\partial\Gamma)} \varepsilon f[x(u_i, v_i), y(u_i, v_i)]\mu\Gamma_i. \tag{46.41}$$

La sumación en estas sumas está extendida a todos los índices i, para los cuales $\overline{\Gamma}_i$ no se interseca con la frontera de Γ. Para la primera suma que figura en el segundo miembro de la igualdad (46.41), se tiene, en virtud de la condicion (46.33) (véase el teorema 5 en el p. 44.3)

$$\lim_{k\to\infty} \sum_{\Gamma_i\in\tau_k(\partial\Gamma)} f[x(u_i, v_i), y(u_i, v_i)]\,|J(u_i, v_i)|\mu\Gamma_i = \iint_{\overline{\Gamma}} f[x(u, v), y(u, v)]\,|J(u, v)|\,dudv.$$

En cuanto a la segunda suma en la igualdad (46.41), ésta tiende a cero cuando $k \to \infty$. En efecto, por ser continua la función $f[x(u, v), y(u, v)]$ en el compacto $\overline{\Gamma}$, es acotada en él, es decir, existe tal constante $C > 0$, que

$$|f[x(u, v), y(u, v)]| \leqslant C, \quad (u, v) \in \overline{\Gamma}.$$

Siendo fijado $\varepsilon_0 > 0$ arbitrario, en virtud de que en $\overline{\Gamma}\varepsilon$ tiende a cero uniformemente para $k \to \infty$, podemos escoger k_0 de modo tal que para $k \geqslant k_0$ se verifique la desigualdad $|\varepsilon| < \dfrac{\varepsilon_0}{c\mu\Gamma}$, cualesquiera que sean $(u_i, v_i) \in \overline{\Gamma}_i$, $\overline{\Gamma}_i \subset \overline{\Gamma}$; entonces

$$\left|\sum_{\Gamma_i\in\tau_k(\partial\Gamma)} \varepsilon f[x(u_i, v_i), y(u_i, v_i)]\mu\Gamma_i\right| \leqslant \sum_{\Gamma_i\in\tau_k(\partial\Gamma)} |\varepsilon|\,|f[x(u_i, v_i), y(u_i, v_i)]|\,\mu\Gamma_i < \frac{\varepsilon_0}{\mu\Gamma} \sum_i \mu\Gamma_i \leqslant \varepsilon_0.$$

Así pues,

$$\lim_{k\to\infty} \sigma_{\tau_k(\partial\Gamma)} = \iint_{\overline{\Gamma}} f[x(u, v), y(u, v)]\,|J(u, v)|\,dudv. \tag{46.42}$$

Precisamente de (46.39) y (46.42) proviene directamente la fórmula (46.32). □

El teorema demostrado se generaliza con facilidad también a un caso algo más general, cuando el jacobiano de la aplicación (46.1) puede reducirse a cero en la frontera de la región de integración, mientras que la propia aplicación falla de ser biunívoca en esta frontera. Con más precisión, es válido el siguiente teorema.

Teorema 2'. *Sean G y G^* unos conjuntos cuadrables abiertos: $G \subset R^2_{uv}$, $G^* \subset R^2_{xy}$, y supongamos que*

$$x = x(u, v),$$
$$y = y(u, v),$$

es una aplicación continua de $\overline{G}$ sobre $\overline{G}^*$ que aplica G sobre G^* de manera biunívoca y continuamente derivable, con la particularidad de que el jacobiano $\dfrac{\partial(x, y)}{\partial(u, v)}$ de dicha aplicación no se reduce a cero en G y es continuamente prolongable a G. En este caso, si la función $f(x, y)$ es continua en el conjunto G^*, se tiene

$$\iint\limits_{G^*} f(x, y)dxdy = \iint\limits_{G} f[x(u, v), y(u, v)]\left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv.$$

DEMOSTRACIÓN. Sea Γ_k, $k = 1, 2, \ldots$ una sucesión de conjuntos abiertos cuadrables acotados cuya frontera se compone de un número finito de curvas suaves a trozos, y

$$\overline{\Gamma}_k \subset G, \quad \Gamma_k \subset \Gamma_{k+1}, \quad \bigcup_{k=1}^{\infty} \Gamma_k = G.$$

En calidad de Γ_k puede tomarse, por ejemplo, una totalidad de todos los puntos interiores del conjunto $S_k(G)$ (véase el p. 44.1). Supongamos que $\Gamma_k^* = F(\Gamma_k)$; entonces Γ_k^* será también un conjunto cuadrable abierto acotado y

$$\overline{\Gamma}_k^* = F(\overline{\Gamma}_k) \subset G^*, \quad \Gamma_k^* \subset \Gamma_{k+1}^*, \quad \bigcup_{k=1}^{\infty} \Gamma_k^* = G^*.$$

Del cumplimiento de estas condiciones se deduce que (véase el teorema 2, p. 31.2)

$$\lim_{k\to\infty} \mu\Gamma_k = \mu G, \quad \lim_{k\to\infty} \mu\Gamma_k^* = \mu G^*. \tag{46.43}$$

Para cada uno de los conjuntos Γ_k, $k = 1, 2, \ldots$, se cumplen todas las condiciones del teorema 2, razón por la cual

$$\iint\limits_{\Gamma_k^*} f(x, y)dxdy = \iint\limits_{\Gamma_k} f[x(u, v), y(u, v)]\left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv. \tag{46.44}$$

La función $f(x, y)$ es continua en $\overline{G}^*$ y por esta razón es integrable en G^*, mientras que la función $f[x(u, v), y(u, v)]\left|\dfrac{\partial(x, y)}{\partial(u, v)}\right|$ es integrable por la misma

razón *) en G. Por ello, como las condiciones (46.43) quedan cumplidas, obtenemos (véase la propiedad 10° en el p. 44.6):

$$\lim_{k\to\infty} \iint_{\bar{G}_k} f(x, y)dxdy = \iint_{\bar{G}^*} f(x, y)dxdy,$$

$$\lim_{k\to\infty} \iint_{\bar{G}_k} f[x(u, v), y(u, v)] \left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv =$$

$$= \iint_{\bar{G}} f[x(u, v), y(u, v)] \left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv. \quad (46.45)$$

Pasando al límite para $k \to \infty$, en la igualdad (46.44) obtendremos, en virtud de la fórmula (46.45), la fórmula buscada para el cambio de las variables en la integral. □

El cambio de las variables en una integral múltiple simplifica a menudo considerablemente la investigación y el cálculo de la integral. Además, a diferencia de una integral sencilla, el objetivo del cambio de variable consiste frecuentemente no en la simplificación de la función subintegral, sino en el paso a una región de integración más simple, incluso a cuenta de la complicación determinada de la función subintegral.

Ejemplo. Calculemos la integral $\iint_{x^2+y^2<1} \cos\pi\sqrt{x^2+y^2}dxdy$.

Con este fin introduzcamos las coordenadas nuevas r, φ (coordenadas polares), rigiéndonos por las fórmulas

$$x = r\cos\varphi, \quad y = r\operatorname{sen}\varphi. \quad (46.46)$$

En este caso $\dfrac{\partial(x, y)}{\partial(r, \varphi)} = \begin{vmatrix} \cos\varphi & -r\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & r\cos\varphi \end{vmatrix} = r$. La aplicación (46.46) aplica el rectángulo $G = \{(r, \varphi) : 0 < r < 1, -\pi < \varphi < \pi\}$ (aquí r y φ se consideran como las coordenadas cartesianas en el plano r, φ), de manera biunívoca, continuamente derivable y con el jacobiano distinto de cero, sobre el círculo $K = \{(x, y) : x^2 + y^2 < 1\}$, del cual se ha excluido el radio dispuesto en la parte negativa del eje Ox, es decir, sobre el conjunto (fig. 197) $G^* = K \setminus \{(x, y) : x \leqslant 0, y = 0\}$.

Entre tanto, el rectángulo cerrado $\bar{G}$ pasa, realizándose la aplicación (46.46), a un círculo cerrado $\bar{G}^* = \bar{K}$, con la particularidad de que en la frontera de $\bar{G}$ esta aplicación ya no es biunívoca. El jacobiano de la aplicación (46.46) es continuo en $\bar{G}$ y, *además*, en un punto de la frontera, a saber, en el origen de coordenadas, se reduce a cero. Todas las condiciones que se imponen sobre la aplicación (46.1) en el teorema 2' de este punto se cumplen para la aplicación (46.46), por lo cual puede aplicarse la fórmula del cambio de variable en la integral

*) Recordemos que, en virtud de las condiciones del teorema, dicha función es continuamente prolongable del conjunto G a $\bar{G}$, con la particularidad de que los valores de la función prolongada en la frontera del conjunto G no influyen en los valores de la integral (véase el p. 44.3).

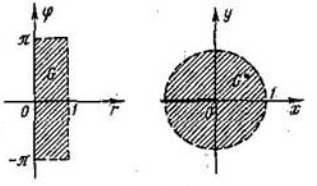


Fig. 197

$$\iint\limits_{x^2+y^2<1} \cos\pi\sqrt{x^2+y^2}\,dxdy = \iint\limits_{\substack{0<r<1\\ -\pi<\varphi<\pi}} r\cos\pi r\,dr\,d\varphi =$$

$$= \int\limits_0^{2\pi} d\varphi \int\limits_0^1 r\cos\pi r\,dr = 2\pi\left[\frac{r\operatorname{sen}\pi r}{\pi}\Big|_0^1 - \frac{1}{\pi}\int\limits_0^1 \operatorname{sen}\pi r\,dr\right] = -\frac{4}{\pi}.$$

La fórmula (46.32) de cambio de variables en la integral doble puede ser demostrada también para el caso más general, en particular, cuando el jacobiano de la aplicación se reduce a cero en la región de integración, mientras que la función subintegral tiene discontinuidades. Si los conjuntos de los puntos indicados tienen medida cero y se aplican también en los conjuntos de medida cero, con la particularidad de que los conjuntos citados dividen las regiones de integración G y G^* en un número finito de conjuntos abiertos, en cada uno de los cuales la función subintegral resulta prolongable de modo que se obtenga una función continua hasta la misma frontera, entonces la fórmula (46.32) se deduce directamente de lo demostrado más arriba.

Ejercicios. Sea $f(x, y) = 2\pi(x^2 - y^2)\operatorname{sen}\pi(x - y)^2$ y sea X un cuadrado cuyos vértices son $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$. Calcúlese la integral $\iint\limits_X f(x, y)dxdy$, sirviéndose del cambio de variables $u = x + y$, $v = x - y$.

2. Pasando a las coordenadas polares, calcúlese la integral $\iint\limits_X f(x, y)dxdy$, donde

a) $f(x, y) = y$, $X = \{(x, y) : x^2 + y^2 < 2x,\ 0 < x < y\}$;

b) $f(x, y) = \sqrt{x^2 + y^2}$, $X = \{(x, y) : x^2 + y^2 < 1,\ x < y < \sqrt{3}x\}$.

46.3. COORDENADAS CURVILÍNEAS

Las fórmulas

$$x = x(u, v),\quad y = y(u, v) \tag{46.47}$$

pueden considerarse no sólo como aplicación, sino también como el paso de un sistema de coordenadas al otro, en el caso general, curvilíneo. Aclaremos, ante todo, la noción de sistema de coordenadas curvilíneas.

Sea G un conjunto abierto en el plano R^2_{xy} y a todo punto $M = (x, y) \in G$, y, por consiguiente, a todo par ordenado de números (x, y), que representan las coordenadas del punto M en el sistema elegido de coordenadas rectangulares, se le ha puesto en correspondencia un par de números (u, v) de un modo tal que a los puntos diferentes M_1 y M_2 corresponden diferentes pares (u_1, v_1) y (u_2, v_2). En este caso suele decirse que en el conjunto G se ha dado el sistema de coordenadas u, v. Si al punto M le corresponde el par (u, v) se escribe $M = (u, v)$. Cualquier par (u, v) es una función del punto $M \in G$, razón por la cual cada uno de sus elementos, u y v, es también una función del punto M: $u = u(M)$, $v = v(M)$, o bien, que es lo mismo, de sus coordenadas cartesianas:

$$\begin{aligned} u &= u(x, y), \\ v &= v(x, y). \end{aligned} \tag{46.48}$$

Viceversa, a todo par (u, v) del conjunto de pares en consideración corresponde un punto $M \in G$, es decir, dicho punto M es una función de los pares (u, v): $M = M(u, v)$, por lo cual sus coordenadas cartesianas x e y son también funciones de los pares citados (u, v). En otras palabras, resultan válidas las fórmulas (46.47) que predeterminan una aplicación, inversa de la aplicación (46.48).

Los conjuntos de puntos $(x, y) \in G$ que satisfacen la condición $u(x, y) = u_0$, y, por consiguiente, $v(x, y) = v_0$, donde u_0 y v_0 son ciertas constantes fijadas, se llaman *líneas coordenadas* en el sistema de coordenadas u, v.

Haciendo uso de la fórmula (46.47), las líneas coordenadas pueden escribirse en la forma

$$\begin{aligned} x &= x(u_0, v), \\ y &= y(u_0, v), \end{aligned} \tag{46.49}$$

y, correspondientemente, en la forma

$$\begin{aligned} x &= x(u, v_0), \\ y &= y(u, v_0). \end{aligned} \tag{46.50}$$

En el caso de las coordenadas cartesianas las líneas coordenadas son rectas; en el caso general, ciertas curvas que se definen por las representaciones (46.49) y (46.50). A esto precisamente se debe la denominación "coordenadas curvilíneas" (fig. 198).

Supondremos que las funciones (46.47) satisfacen en G todas las condiciones, bajo las cuales se ha deducido la fórmula (46.32) de cambio de variable en la integral, en particular, dichas funciones son continuamente derivables y el jacobiano $\dfrac{\partial(x, y)}{\partial(u, v)}$ es distinto de cero en G. Debido a esto, las líneas coordenadas en el entorno de todo punto perteneciente a G son unas curvas continuamente derivables.

Investiguemos qué sentido tendrá en este caso el módulo del jacobiano. Fijemos algunos valores u_0, Δu, v_0, Δv. Supongamos que $M_0 = (u_0, v_0)$, Γ es el conjunto de todos los puntos cuyas coordenadas curvilíneas u, v satisfacen las desigualdades $u_0 < u < u_0 + \Delta u$, $v_0 < v < v_0 + \Delta v$, y sea $\bar{\Gamma} \subset G$. El conjunto Γ se denomina

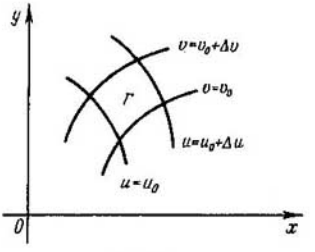


Fig. 198

paralelogramo coordenado (curvilíneo). El conjunto Γ está abierto (¿por qué?) y su frontera representa un contorno suave a trozos (se compone de las curvas del tipo $x = x(u_0, v), y = y(u_0, v)$, donde $v_0 \leqslant v \leqslant v_0 + \Delta v$, etc.) y por eso Γ es una región cuadrable. Calculemos su área (véase la fig. 198). Al aplicar la fórmula de cambio de variable en la integral y el teorema integral del valor medio (véase el p. 44.6), obtendremos

$$\mu\Gamma = \iint_{\Gamma} dxdy = \iint\limits_{\substack{u_0 < u < u_0 + \Delta u \\ v_0 < v < v_0 + \Delta v}} \left|\frac{\partial(x, y)}{\partial(u, v)}\right| dudv =$$

$$= \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M \int_{u_0}^{u_0+\Delta u} du \int_{v_0}^{v_0+\Delta v} dv = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M \Delta u \Delta v, \quad M \in \bar{\Gamma}.$$

Por ser continuamente derivables las funciones (46.47),

$$\left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_{M_0} + \varepsilon,$$

donde $\lim\limits_{\Delta u^2 + \Delta v^2 \to 0} \varepsilon = 0$. De este modo

$$\mu\Gamma = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_{M_0} \Delta u \Delta v + \varepsilon \Delta u \Delta v. \tag{46.51}$$

La fórmula (46.51) señala que el módulo del jacobiano en el punto (u_0, v_0) representa el coeficiente de la parte principal del área del paralelogramo coordenado con vértice en el punto (u_0, v_0) respecto del producto $\Delta u \Delta v$ para $\Delta u^2 + \Delta v^2 \to 0$. Esta observación se utiliza a menudo en la práctica al calcular el jacobiano de la transformación de las coordenadas curvilíneas en las cartesianas. Mostrémoslo en el ejemplo de las coordenadas polares r, φ. Fijemos ciertos valores $r, r + \Delta r, \varphi$ y

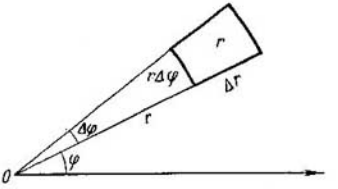


Fig. 199

$\varphi + \Delta\varphi$, y examinemos un paralelogramo coordenado Γ (fig. 199), formado por las líneas coordenadas r, $r + \Delta r$, φ y $\varphi + \Delta\varphi$. Las longitudes de sus dos lados son iguales a Δr, y $r\Delta\varphi$, respectivamente. Al calcular el área de este paralelogramo, como si fuera éste un rectángulo ordinario, tendremos

$$\mu\Gamma = r\Delta r\Delta\varphi.$$

De este modo, el coeficiente del producto $\Delta r\Delta\varphi$ resultó ser igual a r, de donde es natural esperar que $\dfrac{\partial(x, y)}{\partial(u, v)} = r$. Efectivamente, es así (véase el ejemplo en el p. 46.2), lo que se debe a que en nuestros cálculos inexactos del área de Γ se ha admitido un error cuyo orden de infinitud es superior que el producto $\Delta r\Delta\varphi$ para $\Delta r^2 + \Delta\varphi^2 \to 0$. Efectivamente, al calcular $\mu\Gamma$ como una diferencia de las áreas de dos sectores, obtendremos

$$\mu\Gamma = \frac{\Delta\varphi}{2\pi}\pi(r + \Delta r)^2 - \frac{\Delta\varphi}{2\pi}\pi r^2 = r\Delta r\Delta\varphi + \frac{1}{2}\Delta r^2\Delta\varphi.$$

46.4. CAMBIO DE VARIABLES EN UNA INTEGRAL n-MÚLTIPLE

Todo lo expuesto en los puntos anteriores de este párrafo, junto con la demostración, se extiende al caso n-dimensional, por lo cual nos limitaremos sólo a la enunciación de los teoremas correspondientes.

Teorema 3. *Supongamos que $G_x \subset R_x^n$ y $G_t \subset R_t^n$ son unos conjuntos abiertos, $x = F(t) = \{x_i = x_i(t_1, \ldots, t_n),\ i = 1, 2, \ldots, n\}$ es la aplicación biunívoca continuamente derivable de G_t sobre G_x, cuyo jacobiano $J(t) = \dfrac{\partial(x_1, \ldots, x_n)}{\partial(t_1, \ldots, t_n)}$ es distinto de cero en G_t. Sea, además, S un cubo n-dimensional:*

$$S = \{t : t_i^{(0)} \leqslant t_i \leqslant t_i^{(0)} + h,\ i = 1, 2, \ldots, n\} \subset G_t,\ t^{(0)} = (t_1^{(0)}, \ldots, t_n^{(0)}).$$

Entonces $\lim\limits_{h \to 0} \dfrac{\mu F(S)}{\mu S} = |J(t^{(0)})|$; *con lo cual, si*

$$\frac{\mu F(S)}{\mu S} = |J(t^{(0)})| + \varepsilon(t^{(0)}, h),$$

para cualquier compacto $A \subset G_t$ la función $\varepsilon(t^{(0)}, h)$, $t^{(0)} \in A$, tiende uniformemente hacia cero en A, cuando $h \to 0$.

Teorema 4. *Sean:*

1) G_x y G_t unos conjuntos abiertos medibles $G_x \subset R^n_x$, $G_t \subset R^n_t$;

2) $x = F(t)$, *la aplicación continua de* $\overline{G}_t$ *sobre* $\overline{G}_x$ *que aplica* G_t *sobre* G_x *de una manera biunívoca y continuamente derivable;*

3) *el jacobiano* $J(t) = \dfrac{\partial(x_1, \ldots, x_n)}{\partial(t_1, \ldots, t_n)}$ *de esta aplicación no se reduce a cero en* G_t *y es continuamente prolongable a* $\overline{G}_t$.

En este caso, si la función $f(x)$ *es continua en* $\overline{G}_x$, *se tiene*

$$\int f(x)\,dG_x = \int f(x(t))\,|J(t)|\,dG_t.$$

Ejercicios. Escríbanse las fórmulas del cambio de variables en las integrales triples para la transformación de coordenadas:

3. $x = r\cos\psi\cos\varphi$, $y = \cos\psi\operatorname{sen}\varphi$, $z = r\operatorname{sen}\psi$, $0 \leqslant r < +\infty$, $0 \leqslant \varphi \leqslant 2\pi$, $-\dfrac{\pi}{2} \leqslant \psi \leqslant \dfrac{\pi}{2}$ (coordenadas esféricas).

4. $x = r\cos\psi$, $y = r\operatorname{sen}\varphi$, $z = z$, $0 \leqslant r < +\infty$, $0 \leqslant \varphi \leqslant 2\pi$, $-\infty < z < +\infty$ (coordenadas cilíndricas).

OBSERVACIÓN. Siendo vigente la fórmula del cambio de variable para cualquier conjunto medible $G \subset R^n$ y cualesquiera coordenadas curvilíneas $u_1, \ldots, u_n$, resulta lícita la siguiente fórmula

$$\int_G \cdots \int dx_1 \ldots dx_n = \int_G \cdots \int \left|\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)}\right| du_1 \ldots du_n.$$

En particular, si

$$\left|\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)}\right| = 1, \tag{46.52}$$

entonces

$$\int_G \cdots \int dx_1 \ldots dx_n = \int_G \cdots \int du_1 \ldots du_n. \tag{46.53}$$

La condición (46.52) se cumple a ciencia cierta, si $u_1, \ldots, u_n$ son también las coordenadas cartesianas en el espacio R^n y, por tanto, se expresan mediante $x_1, \ldots, x_n$ con ayuda de una transformación lineal cuyo determinante es igual a ± 1;

$$x_i = \sum_{j=1}^{n} c_{ij}u_j, \quad i = 1, 2, \ldots, n, \quad \det\|c_{ij}\| = \pm 1.$$

El primer miembro de la fórmula (46.53) es igual a la medida μG del conjunto G "en las coordenadas $x_1, \ldots, x_n$" (véase la propiedad 1° de la integral múltiple en el p. 44.6) y el segundo miembro de la misma, cuando $u_1, \ldots, u_n$ son las coordenadas cartesianas, es igual, correspondientemente, a la medida del conjnto G "en las coordenadas $u_1, \ldots, u_n$". De este modo, la fórmula (46.53) indica que la medida de un conjunto abierto medible no depende de cómo se elige el sistema de coordenadas cartesjanas.

A decir verdad, debemos observar que al demostrar la fórmula del cambio de variables en la integral múltiple, nos hemos servido de aquel hecho demostrado en la geometría que, realizándose las aplicaciones afines (es decir, aplicaciones lineales regulares), el valor absoluto del determinante de la transformación, es igual a la razón entre el volumen del paralelepípedo, que es la imagen de cierto cubo, y el volumen del último.

Ejercicios. 5. Sea $G = \{(x, y, z) : 1 < x < 2;\ 1 < x + y < 3;$ $1 < x + y + z < 5\}$. Calcúlese la integral $\iiint_G \frac{dxdydz}{(x+y)(x+y+z)}$, pasando a las variables u, v, w que están ligadas con x, y, z mediante las correlaciones $x + y + z = u$, $x + y = uv$, $x = uvw$.

6. Sea $G = \{(x, y, z) : x < yz < 2x;\ y < zx < 2y;\ z < xy < 2z\}$. Calcúlese la integral $\iiint_G xyz\,dxdydz$, pasando a las variables u, v, w que están ligadas con x, y, z por las correlaciones $ux = yz$, $vy = zx$, $wz = xy$.

§ 47. INTEGRALES CURVILÍNEAS

47.1. INTEGRALES CURVILÍNEAS DE PRIMERA ESPECIE

Sea dada en el espacio tridimensional R^3 una curva $\gamma = \{r(t);\ a \leqslant t \leqslant b\}$ (véase § 16). Consideraremos las funciones unívocas F definidas en los puntos $r(t)$ de esta curva: $F = F(r(t))$. Si $\rho(t)$, $\alpha \leqslant \tau \leqslant \beta$, es alguna otra representación de la misma curva γ y si $t = t(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, es la aplicación del segmento $[\alpha, \beta]$ sobre el segmento $[a, b]$ que realiza la equivalencia de dichas representaciones (es decir, si $t = t(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, es la transformación admisible del parámetro, véase el p. 16.1), entonces, como el valor de la función F se determina sólo por un punto de la curva, tendremos

$$F(\rho(\tau)) = F(r(t)), \quad t = (\tau), \quad \alpha \leqslant \tau \leqslant \beta.$$

Las funciones en consideración F admiten, en el caso general, diferentes valores en los puntos de la curva, correspondientes a los valores distintos del parámetro, pero coincidentes como puntos de un espacio (véanse los puntos múltiples en el p. 16.1). Este punto de vista corresponde a la interpretación física de la curva γ, por ejemlo, cuando ésta representa una trayectoria que sigue un punto material en movimiento, y a la de la función F, cuando F interviene como una fuerza que actúa contra dicho punto y que depende no sólo de la posición del punto en el espacio, sino también del instante en que este punto se encuentra en el lugar dado. Además, semejante enfoque proporciona también algunas ventajas matemáticas que se pondrán de manifiesto en lo que sigue.

De lo expuesto se desprende que las funciones citadas, definidas en una curva, no pueden considerarse como funciones prefijadas en cierto conjunto del espacio R^3, por lo cual, en palabras estrictas, no pueden designarse mediante $F(x, y, z)$, donde x, y, z son las coordenadas cartesianas de los puntos espaciales. No obstante

en los problemas que se considerarán abajo tal designación lleva un carácter tradicional y por esta razón se empleará en nuestros razonamientos. Si se recuerda siempre que en dichos problemas se trata de las funciones, definidas en los puntos de las curvas, la designación citada no nos llevará a las equivocaciones.

Supongamos ahora que está dada una curva orientada rectificable γ, con la particularidad de que $r(s) = [x(s), y(s), z(s)]$; $0 \leqslant s \leqslant S]$ es su representación en la cual a título de parámetro interviene la longitud variable del arco s, y sean $A = r(0)$ y $B = r(S)$ los puntos inicial y final de esta curva. En este caso escribiremos $\gamma = \overset{\frown}{AB}$. La curva de orientación opuesta se designará con $\overset{\frown}{BA}$.

Definición 1. *Supongamos que en los puntos $r(s)$ de la curva γ se ha dado cierta función F. La expresión* $\int\limits_{\overset{\frown}{AB}} F(x, y, z)ds$ *que se determina según la fórmula*

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)ds = \int\limits_{0}^{S} F(x(s), y(s), z(s))ds, \tag{47.1}$$

se denomina integral curvilínea de primera especie de la función F a lo largo de la curva $\overset{\frown}{AB}$.

Dicha integral se designa también mediante los símbolos

$$\int\limits_{\overset{\frown}{AB}} F[r(s)]ds \quad \text{y} \quad \int\limits_{\gamma} F[r(s)]ds, \text{ o bien, más brevemente, } \int\limits_{\gamma} Fds.$$

De este modo, aunque la definición de integral curvilínea de primera especie está relacionada con el concepto de curva, es decir, con una imagen geométrica, ella se reduce a una integral ordinaria por un segmento, por lo cual a la integral curvilínea se extienden todas las propiedades de la integral ordinaria.

Demos a conocer algunas propiedades específicas de la integral (47.1).

1°. $\int\limits_{\overset{\frown}{AB}} ds = S.$

Esto es obvio.

2°. *Si la función F es continua en los puntos de la curva γ como función del parámetro s, es decir, si es continua la función $F[r(s)]$, $0 \leqslant s \leqslant S$, entonces* $\int\limits_{\gamma} Fds$ *existe.*

En efecto, de acuerdo con la definición (47.1), la integral $\int\limits_{\gamma} Fds$ se reduce a la integral $\int\limits_{0}^{S} F[x(s), y(s), z(s)]ds$ de una función continua por un segmento que, como se sabe, existe.

3°. *Una integral curvilínea de primera especie no depende de la orientación de la curva:*

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)ds = \int\limits_{\overset{\frown}{BA}} F(x, y, z)ds.$$

Efectivamente, supongamos que $M = r(s)$ es un punto de la curva $\overset{\frown}{AB}$, y s es la longitud del arco $\overset{\frown}{AM}$. Si $\sigma = S - s$, entonces σ es igual a la longitud del arco $\overset{\frown}{BM}$

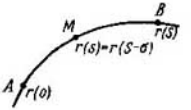


Fig. 200

(fig. 200). La función $r = r(S - \sigma)$; $0 \leqslant \sigma \leqslant S$, es la representación de la curva $\overset{\frown}{BA}$, por lo cual, al realizar en la integral (47.1) el cambio de variable $s = S - \sigma$, y al notar que $ds = -d\sigma$, obtenemos

$$\int_{\overset{\frown}{AB}} F(x, y, z)ds = \int_0^S F[x(s), y(s), z(s)]ds =$$

$$= -\int_S^0 F[x(S-\sigma), y(S-\sigma), z(S-\sigma)]d\sigma =$$

$$= \int_0^S F[x(S-\sigma), y(S-\sigma), z(S-\sigma)]d\sigma = \int_{\overset{\frown}{BA}} F(x, y, z)d\sigma.$$

Esta propiedad de la integral curvilínea de primera especie se debe a que la longitud del arco de una curva, de acuerdo con la definición, se considera positiva, independientemente del extremo a partir del cual dicha longitud se calcula.

Antes de pasar a la siguiente propiedad, hemos de notar que $\int_\gamma Fds$, al igual que cualquier otra integral, es un límite de las sumas integrales correspondientes; la peculiaridad del caso dado consiste sólo en que estas sumas pueden ser descritas en términos geométricos ligados con la curva γ, a lo largo de la cual se realiza la integración. He aquí una formulación más precisa.

4°. *Sea* $\tau = \{s_i\}_{i=0}^{i_0}$ *la partición del segmento* $[0, S]$, $\xi_i \in [s_{i-1}, s_i]$; $\Delta s_i = s_i - s_{i-1}$, *la longitud del arco de la curva* γ *entre los puntos* $r(s_{i-1})$ y $r(s_i)$, $i = 1, 2, \ldots, i_0$, *mientras que* $\sigma_\tau = \sum_{i=1}^{i_0} F[r(\xi_i)]\Delta s_i$. *Entonces, si la función* $F[r(s)]$ *es integrable según Riemann en el segmento* $[0, S]$, *se tiene*

$$\lim_{\sigma_\tau \to 0} \sigma_\tau = \int_\gamma Fds. \tag{47.2}$$

Efectivamente, σ_τ es, evidentemente, la suma integral de Riemann de la integral $\int_0^S F[r(s)]ds$, y, por ende, la fórmula (47.2) proviene directamente de la (47.1).

La fórmula (47.1) es muy cómoda en el estudio de las propiedades de la integral $\int_\gamma Fds$, no obstante, no es así ni mucho menos en el cálculo de la misma, puesto

que suele ser muy difícil o incluso prácticamente imposible encontrar una representación de la curva dada, en la que por parámetro se toma la longitud variable del arco. Por esta razón indiquemos la fórmula de la integral $\int_\gamma Fds$ para cualquier representación paramétrica de la curva γ.

5°. *Supongamos que γ es una curva suave* (véase la definición 16 en el p. 16.4), $r(t) = [\varphi(t), \psi(t), \chi(t); a \leqslant t \leqslant b]$ *es su representación continuamente derivable y, por consiguiente,* $\varphi'^2(t) + \psi'^2(t) + \chi'^2(t) > 0$, $a \leqslant t \leqslant b$ *).

Sea la función F continua en la curva γ (en aquel sentido que la función $F[r(t)]$ es continua en el segmento $[a, b]$). En este caso

$$\int_\gamma F(x, y, z)ds = \int_a^b F[\varphi(t), \psi(t)\chi(t)]\sqrt{\varphi'^2(t)+\psi'^2(t)+\chi'^2(t)}dt. \qquad (47.3)$$

En efecto, bajo las suposiciones asumidas, la curva γ es rectificable y la longitud variable del arco $s = s(t)$ puede tomarse como parámetro (véase el corolario 2 del teorema 2 en el p. 16.5), a consecuencia de lo cual la integral $\int_\gamma Fds$ tiene sentido.

Al realizar el cambio de variable $s = s(t)$ en el segundo miembro de la igualdad (47.1) y recordar que (véase el p. 16.5),

$$\frac{ds}{dt} = \sqrt{x'^2_t + y'^2_t + z'^2_t},$$

obtendremos la fórmula (47.3).

De (47.3) se infiere que para la curva dada el valor de la integral que figura en el segundo miembro de la igualdad (47.3) no depende de cómo se elige el parámetro en la curva, pues, cualquiera que sea la elección del parámetro, dicha integral es igual a la integral que figura en el primer miembro de esta igualdad.

47.2. INTEGRALES CURVILÍNEAS DE SEGUNDA ESPECIE

Existe una serie de problemas matemáticos y aplicados que conducen a las integrales curvilíneas de otro género. Por ejemplo, si $r = r(t)$ es el radio vector de un punto material en movimiento y $F = F(t)$ es la fuerza que actúa contra este punto, entonces sería natural determinar el trabajo de la fuerza F a lo largo de la trayectoria Γ del punto como una integral $\int_\Gamma Fdr$, o bien, siempre que $F = (P, Q, R)$, y $dr = (dx, dy, dz)$, en la inscripción de coordenadas como una integral

$$\int_\Gamma Pdx + Qdy + Rdz. \qquad (47.4)$$

*) Recordemos que esta condición es indicio de que la curva está privada de puntos singulares (véase la definición 15 en el p. 16.4).

Recordando que (véase el p. 16.5)

$$\frac{dx}{ds} = \cos\varphi, \quad \frac{dy}{ds} = \cos\beta, \quad \frac{dz}{ds} = \cos\gamma, \tag{47.5}$$

donde $t = (\cos\varphi, \cos\beta, \cos\gamma)$ es el vector unidad tangente, la integral (47.4) puede ser representada formalmente en la forma

$$\int_{\Gamma} (P\cos\alpha + Q\cos\beta + R\cos\gamma)ds.$$

Enunciemos ahora la definición estricta de las integrales del tipo (47.4). Sea $\gamma = \overset{\frown}{AB}$ una curva orientada suave, es decir, una curva orientada continuamente derivable, privada de puntos singulares. En este caso existe su representación continuamente derivable

$$r(t) = \{x = \varphi(t), y = \psi(t), z = \chi(t); a \leqslant t \leqslant b\}, \quad A = r(a), \quad B = r(b),$$

tal que

$$\varphi'^2(t) + \psi'^2(t) + \chi'^2(t) > 0, \quad a \leqslant t \leqslant b.$$

Sea $s = s(t)$ la longitud variable del arco, $0 \leqslant s \leqslant S$, S es la longitud de toda la curva γ, calculada a partir del extremo A, $(\cos\alpha, \cos\beta, \cos\gamma)$ es el vector unidad tangente a la curva, $\alpha = \alpha(s), \beta = \beta(s), \gamma = \gamma(s), 0 \leqslant s \leqslant S$, y supongamos que la función F, al igual que en el punto anterior, está definida en el conjunto $\{r(t), a \leqslant t \leqslant b\}$ de todos los puntos de la curva γ.

Definición 2. *La integral* $\int_{\overset{\frown}{AB}} F(x, y, z)dx$ *se determina según la fórmula*

$$\int_{\overset{\frown}{AB}} F(x, y, z)dx = \int_{\overset{\frown}{AB}} F(x, y, z)\cos\alpha ds. \tag{47.6}$$

Análogamente, según la definición, debe ser

$$\int_{\overset{\frown}{AB}} F(x, y, z)dy = \int_{\overset{\frown}{AB}} F(x, y, z)\cos\beta ds,$$

$$\int_{\overset{\frown}{AB}} F(x, y, z)dz = \int_{\overset{\frown}{AB}} F(x, y, z)\cos\gamma ds, \tag{47.7}$$

Las integrales del tipo (47.6) *y* (47.7) *se llaman integrales curvilíneas de segunda especie de la función F a lo largo de la curva $\overset{\frown}{AB}$.*

El carácter natural de estas definiciones se ve de las fórmulas (47.5).

Demos a conocer algunas propiedades de las integrales curvilíneas de segunda especie, limitándonos, para brevedad, sólo al caso de la integral (47.6).

1°. *Si la función F es continua en la curva γ, es decir, si es continua la función $F[r(t)]$, $a \leqslant t \leqslant b$, la integral (47.6) existe.*

En efecto, admitidas las suposiciones respecto de la curva γ, la función $t = t(s)$ (t es el parámetro en la curva γ, y s es la longitud variable del arco) es continuamente derivable en el segmento $[0, S]$, por lo cual la función $\cos\alpha = \dfrac{dx}{dt}\dfrac{dt}{ds}$ es continua

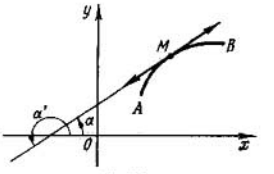


Fig. 201

en este segmento y, por lo tanto, en virtud de la propiedad 2° de las integrales curvilíneas de primera especie (véase el p. 47.1), la integral (47.6) existe. □

En lo que sigue en este punto supondremos, para simplificar, que la función F es continua en la curva γ. En este caso todas las integrales que van abajo existen a ciencia cierta.

2°. *Una integral curvilínea de segunda especie cambia de signo cuando la curva varía su orientación, es decir,*

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)dx = -\int\limits_{\overset{\frown}{BA}} F(x, y, z)dx.$$

En efecto, si α es el ángulo formado por el sentido positivo de una tangente a la curva $\overset{\frown}{AB}$ y el eje Ox, y α', el ángulo formado por el sentido positivo de la tangente a la curva $\overset{\frown}{BA}$ y el eje Ox, entonces para los puntos correspondientes tendremos $\alpha' = \alpha + \pi$ (fig. 201), y, por lo tanto, $\cos\alpha' = -\cos\alpha$.

Ahora, haciendo uso de la propiedad de independencia de una integral curvilínea de primera especie de la orientación de la curva (véase el p. 47.1), obtendremos

$$\int\limits_{\overset{\frown}{BA}} F(x, y, z)dx = \int\limits_{\overset{\frown}{BA}} F(x, y, z)\cos\alpha' ds = -\int\limits_{\overset{\frown}{BA}} F(x, y, z)\cos\alpha ds =$$

$$= -\int\limits_{\overset{\frown}{AB}} F(x, y, z)\cos\alpha ds = -\int\limits_{\overset{\frown}{AB}} F(x, y, z)dx.$$

De este modo, dicha propiedad que posee la integral curvilínea de segunda especie se desprende del hecho de que las integrales curvilíneas

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)dx \quad \text{y} \quad \int\limits_{\overset{\frown}{BA}} F(x, y, z)dx$$

son iguales a las integrales curvilíneas correspondientes de primera especie, cuyas expresiones subintegrales se diferencian sólo en signo. □

3°. *Si F es una función continua en la curva γ, para la integral (47.6) será válida la fórmula*

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)dx = \int\limits_{a}^{b} F[\varphi(t), \psi(t), \chi(t)]\varphi'(t)dt. \tag{47.8}$$

En efecto, de acuerdo con la definición (47.6),

$$\int\limits_{\overset{\frown}{AB}} F(x, y, z)ds = \int\limits_{0}^{S} F[x(s), y(s), z(s)]\cos\alpha(s)ds.$$

Al realizar el cambio de la variable $s=s(t)$ en la integral que figura en el segundo miembro de esta igualdad y notar que (véase (47.5)) $\cos\alpha=\frac{dx}{ds}=\frac{x'_t}{s'_t}$, obtendremos

$$\int_0^S F[x(s), y(s), z(s)]\cos\alpha(s)ds =$$

$$= \int_a^b F[\varphi(t), \psi(t), \chi(t)]\frac{x'_t}{s'_t}s'_t dt =$$

$$= \int_a^b F[\varphi(t), \psi(t), \chi(t)]\varphi'_t(t)dt \quad \square$$

Ha de notarse que hemos demostrado a la vez que la integral en el segundo miembro de esta fórmula no depende de la elección del parámetro en la curva el que conserva la orientación de ésta.

En el caso particular cuando como parámetro t puede tomarse la variable x, es decir, cuando la curva γ posee la representación $y = y(x)$, $z = z(x)$, $a \leqslant x \leqslant b$, y, por lo tanto, está privada de puntos múltiples, la función F será función unívoca no sólo de los puntos de la curva, sino también de los puntos correspondientes del espacio (en este caso a diferentes puntos de la curva les corresponden puntos distintos del espacio y viceversa).

La fórmula (47.8) admite en este caso la siguiente forma

$$\int_\gamma F[x, y, z)dx = \int_a^b F[(x, y(x), z(x)]dx. \tag{47.9}$$

4°. *La integral* $\int_{\overset{\frown}{AB}} F(x, y, z)dx$ *es el límite de las sumas integrales correspondientes que se describen en los términos ligados con la curva* γ, *con más precisión: sea* $\tau = \{t_i\}_{i=0}^{i=i_0}$ *la partición del segmento* $[a, b]$, *y* δ_τ, *su finura,* $\xi_i \in [t_{i-1}, t_i]$, $i = 1, 2, \ldots, i_0$, *y*

$$\tilde{\sigma}_\tau = \sum_{i=1}^{i_0} F[r(\xi_i)]\Delta x_i,$$

donde $\Delta x_i = x(t_i) - x(t_{i-1})$, *entonces*

$$\lim_{\delta_\tau \to 0} \tilde{\sigma}_\tau = \int_{\overset{\frown}{AB}} F(x, y, z)dx. \tag{47.10}$$

En efecto, según la fórmula de los incrementos finitos de Lagrange, $\Delta x_i = \varphi'(\eta_i)\Delta t_i$, donde

$$\eta_i \in [t_{i-1}, t_i], \quad \Delta t_i = t_i - t_{i-1}, \quad i = 1, 2, \ldots, i_0,$$

por lo cual

$$\tilde{\sigma}_\tau = \sum_{i=1}^{i_0} F[r(\xi_i)]\varphi'(\eta_i)\Delta t_i.$$

Pongamos

$$\sigma_\tau = \sum_{i=1}^{i_0} F[r(\xi_i)]\varphi'(\xi_i)\Delta t_i.$$

La suma σ_τ es, para la función $F[r(t)]\varphi(t)$, una suma integral de Riemann, razón por la cual

$$\lim_{\delta_\tau \to 0} \sigma_\tau = \int_a^b F[r(t)]\varphi'(t)dt. \tag{47.11}$$

Por otra parte,

$$|\tilde{\sigma}_\tau - \sigma_\tau| \leqslant \sum_{i=1}^{i_0} |F[r(\xi_i)]|\,|\varphi'(\eta_i) - \varphi'(\xi_i)|\Delta t_i \leqslant$$

$$\leqslant \omega(\delta_\tau; \varphi')(b-a) \sup_{a \leqslant t \leqslant b} |F[r(t)]|,$$

donde $\omega(\delta; \varphi')$ es el módulo de continuidad de la funcion φ'. Se sabe que de la continuidad de la función $F[r(t)]$ en el segmento $[a, b]$ proviene que $\sup_{a \leqslant t \leqslant b} |F[r(t)]| < \infty$, y de la continuidad de la función φ' en el mismo segmento se deduce que $\lim_{\delta_\tau \to 0} \omega(\delta_\tau; \varphi') = 0$, por esta razón se tiene $\lim_{\delta_\tau \to 0} (\tilde{\sigma}_\tau - \sigma_\tau) = 0$. Teniéndolo en cuenta, en virtud de (47.11), obtenemos

$$\lim_{\delta_\tau \to 0} \sigma_\tau = \int_a^b F[r(t)]\varphi'(t)dt.$$

De aquí, de acuerdo con la propiedad 3°, se deduce la fórmula (47.10). □

Hemos indicado sólo aquellas propiedades de las integrales curvilíneas que están ligadas con el carácter específico de su definición y con la curva a lo largo de la cual se realiza la integración. Es natural, pues, que, por cuanto las integrales en consideración se reducen a las integrales ordinarias en un segmento, a las primeras se extienden también las diversas propiedades propias para las segundas (linealidad respecto de las funciones a integrar, teorema integral del valor medio, etc.).

47.3. AMPLIACIÓN DE LA CLASE DE TRANSFORMACIONES ADMISIBLES DEL PARÁMETRO DE UNA CURVA

Una curva orientada continuamente derivable y privada de puntos singulares se ha definido (véase el p. 16.1 y 16.2*) como aquella que dispone de tales representaciones vectoriales continuamente derivables $r(t)$, $a \leqslant t \leqslant b$ que $r'(t) \neq 0$ en el segmento $[a, b]$. A título de transformaciones admisibles del parámetro se consideraban las funciones siguientes

$$t = t(\tau), \quad \alpha \leqslant \tau \leqslant \beta, \quad t(\alpha) = a, \quad t(\beta) = b,$$

que eran continuamente derivables y tenían derivada positiva en el segmento $[a, b]$. No obstante, este requisito resulta, a veces, demasiado pesado. Por ejemplo, para un arco γ de la circunferencia unidad con centro en el origen de coordenadas las representaciones

$$y = \sqrt{1 - x^2}, \quad 0 \leqslant x \leqslant 1, \quad \text{y} \quad x = \operatorname{sen} t, \quad y = \cos t, \quad 0 \leqslant t \leqslant \frac{\pi}{2},$$

resultan ser no equivalentes en este sentido. Además, la misma representación $y = \sqrt{1 - x^2}$, $0 \leqslant x \leqslant 1$, no determina en nuestro sentido la curva continuamente derivable, puesto que para $x = 1$ ella no tiene derivada. Es por eso que sería natural ampliar la clase de transformaciones admisibles de los parámetros y de representaciones admisibles de las curvas continuamente derivables. Podemos hacerlo del modo siguiente.

Consideraremos una totalidad de las representaciones vectoriales $r = r(t)$, $a \leqslant t \leqslant b$, continuas en el segmento $[a, b]$ y continuamente derivables en el intervalo (a, b). Se llamará *transformación admisible del parámetro* toda función $t = t(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, $t(\alpha) = a$, $t(\beta) = b$, continua en el segmento $[\alpha, \beta]$, continuamente derivable y que tiene en el intervalo (α, β) una derivada positiva. Como hasta ahora, dos representaciones se llamarán *equivalentes*, si se puede pasar de una a la otra con ayuda de una transformación admisible del parámetro.

Definición 3. *Una clase de representaciones equivalentes del tipo indicado prefija una curva continuamente derivable, si en la clase citada existe por lo menos una representación $r = r(t)$, $a \leqslant t \leqslant b$, que es continuamente derivable en todo el segmento $[a, b]$.*

Definición 4. *Una curva continuamente derivable lleva el nombre de curva sin puntos singulares o, en la forma más breve, curva suave, si para cierta representación suya $r(t)$, $a \leqslant t \leqslant b$ (y, por lo tanto, para todas las representaciones suyas) se cumple la condición $r'(t) \neq 0$, $a < t < b$.*

En el sentido de esta definición dos representaciones citadas del arco de una circunferencia resultan ser equivalentes y definen una curva suave.

Quedan en vigor también todas las definiciones aducidas anteriormente de las integrales curvilíneas y sus propiedades, si se toma, por supuesto, en consideración el hecho de que para ciertas representaciones de las curvas puede obtenerse una integral impropia.

Cabe subrayar que la ampliación de la clase de representaciones de una curva permite *calcular* la integral curvilínea para las más diversas representaciones de la curva. Por ejemplo, la integral $\int_\gamma P(x, y)dy$, donde γ es el arco de la circunferencia unidad considerado más arriba y P, una función continua en γ, puede calcularse sirviéndose de ambas representaciones mencionadas:

$$\int_\gamma P(x, y)dy = -\int_0^1 P(x, \sqrt{1 - x^2}) \frac{x\,dx}{\sqrt{1 - x^2}},$$

$$\int_{\gamma} P(x, y)dy = -\int_{0}^{\pi/2} P(\operatorname{sen} t, \cos t)\operatorname{sen} t\, dt.$$

En el primer caso aquí puede obtenerse una integral impropia.

Además de esto, *al demostrar* los teoremas pueden elegirse las "representaciones buenas", es decir, aquellas que son continuamente derivables hasta los extremos del segmento, y los razonamientos realizados resultarán válidos también para el concepto ampliado de curva.

Ejercicio 1. Demuéstrese que para la nueva definición de la curva continuamente derivable $\gamma = \{x(t), y(t), z(t)\}$ su longitud se expresa mediante la fórmula $\int_a^b \sqrt{x'^2 + y'^2 + z'^2}\, dt$, donde la integral escrita es, en el caso general, impropia.

47.4. INTEGRALES CURVILÍNEAS A LO LARGO DE LAS CURVAS SUAVES A TROZOS

Definición 5. *Si la curva γ es suave a trozos, es decir, puede ser representada como reunión de un número finito de las curvas suaves $\gamma_1, \gamma_2, \ldots, \gamma_k$, y la función $F(x, y, z)$ está definida, como antes, en los puntos de la curva γ, entonces, por definición, pongamos*

$$\int_{\gamma} F(x, y, z)dx = \sum_{i=1}^{k} \int_{\gamma_i} F(x, y, z)dx.$$

Si γ es una curva suave a trozos y $x = x(t), y = y(t), z = z(t), a \leqslant t \leqslant b$, es su representación suave a trozos, se escribirá también

$$\int_{\gamma} F(x, y, z)dx = \int_{a}^{b} F[x(t), y(t), z(t)]x'(t)dt$$

(aquí la derivada $x'(t)$ puede ser no definida en el número finito de los puntos del segmento $[a, b]$), con la particularidad de que la integral que figura en el segundo miembro de la igualdad se entiende, por regla general, en el sentido impropio.

Las definiciones análogas tienen lugar también para las integrales del tipo (47.7). En adelante tropezaremos con las sumas de las integrales del tipo (47.6) y (47.7), es decir, con las integrales del tipo (47.4), donde P, Q y R son ciertas funciones definidas en los puntos de la curva γ. De acuerdo con las definiciones (47.6) y (47.7), es lícita la fórmula

$$\int_{\gamma} P dx + Q dy + R dz = \int_{\gamma} (P\cos\alpha + Q\cos\beta + R\cos\gamma)ds.$$

OBSERVACIÓN 1. Si Γ denota una totalidad finita de curvas orientadas suaves a trozos γ_i, $i = 1, 2, \ldots, k$, entonces, según la definición,

$$\int_{\Gamma} Fds = \sum_{i=1}^{k} \int_{\gamma_i} Fds, \quad \int_{\Gamma} Fdx = \sum_{i=1}^{k} \int_{\gamma_i} Fdx, \text{ etc.}$$

OBSERVACIÓN 2. Hemos definido las integrales curvilíneas para las curvas dispuestas en el espacio tridimensional R^3. De un modo enteramente análogo se definen dichas integrales también para las curvas que se disponen en cualquier espacio n-dimensional R^n, $n = 2, 3, 4, \ldots$, . Las integrales curvilíneas en un espacio n-dimensional poseen las propiedades que son análogas a las consideradas en el caso tridimensional y sus demostraciones son también las mismas que se han realizado anteriormente. Por eso no nos detendremos en las formulaciones ni tampoco en las demostraciones de las afirmaciones correspondientes.

Ejercicio 2. Demuéstrese que las definiciones, aducidas en el presente punto, de las integrales cirvilíneas a lo largo de las curvas suaves a trozos no dependen de cómo se dividen dichas curvas en arcos suaves.

3. Calcúlese la integral curvilínea $\int_{\Gamma} \sqrt{x+2y}dx + \sqrt{x+y}dy$, donde Γ es un contorno triangular cuyos vértices son $O(0; 0)$, $B(2; 0)$, $C(2; 4)$.

4. Calcúlese la integral curvilínea $\int_{\Gamma} \frac{x^2dy - y^2dx}{x^{5/3} + y^{5/3}}$, donde Γ es el arco de una astroide $x = a\cos^3 t$, $y = a\operatorname{sen}^3 t$, limitada con los puntos $(a, 0)$ y $(0, a)$.

47.5. FÓRMULA DE GREEN

Definición 6. *Supongamos que un contorno sencillo cerrado γ es la frontera de la región plana y acotada G. Si la orientación del contorno es tal que durante el recorrido del contorno γ, correspondiente al crecimiento del parámetro, la región G queda a la izquierda (tal recorrido se llama corrientemente recorrido del contorno en dirección contraria al movimiento de las agujas de un reloj), dicha orientación se denomina positiva; en el caso contrario (es decir, cuando el recorrido del contorno se realiza en el sentido de las agujas de un reloj), se denomina negativa* (fig. 202).

Un contorno orientado positivo se designará γ^+, y un contorno orientado negativo, γ^-. Estas nociones están definidas no de modo estricto y no en términos matemáticos exactos. No obstante, nos abstengamos de dar aquí las definiciones precisas, ya que, por una parte, esto no podría hacerse sucintamente y, por otra parte, en lo que sigue la orientación se indicará cada vez concretamente. De este modo, nuestra definición "general" de la orientación positiva y negativa de un contorno sencillo cerrado nos servirá sólo para los fines de la claridad geométrica de los problemas que se considerarán más abajo.

En adelante la región plana G, cuya clausura puede representarse en la forma (45.1) y (45.13) simultáneamente (véase la fig. 168), se llamará, para abreviar, *región elemental.*

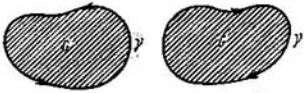


Fig. 202

Teorema 1 (formula de Green *). *Sea G una región plana y supongamos que su frontera γ es un contorno suave a trozos. Supongamos también que la región G* puede ser partida en un número finito de las regiones elementales G_i **) que tienen *fronteras suaves a trozos γ_i, $i = 1, 2, \ldots, k$. Luego, sean dadas en la región cerrada $\overline{G}$ las funciones $P(x, y)$ y $Q(x, y)$ que son continuas en $\overline{G}$ junto con sus derivadas parciales $\dfrac{\partial P}{\partial y}$ y $\dfrac{\partial Q}{\partial x}$.* ***) *Bajo las condiciones mencionadas resulta válida la fórmula*

$$\iint\limits_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \int\limits_{\gamma+} P dx + Q dy. \tag{47.12}$$

DEMOSTRACIÓN. Al principio supongamos que la región G es elemental de por sí y, por consiguiente, su frontera puede ser representada como reunión de las gráficas de dos funciones continuamente derivables a trozos $\varphi(x)$ y $\psi(x)$, $\varphi(x) \leqslant \psi(x)$, $a \leqslant x \leqslant b$, y, quizás, de los segmentos de las rectas $x = a$ y $x = b$, y también como reunión de dos gráficas de las funciones continuamente derivables a trozos $\alpha(y)$ y $\beta(y)$, $\alpha(y) \leqslant \beta(y)$, $c \leqslant y \leqslant d$, y, quizás, de los segmentos de las rectas $y = c$ e $y = d$ (fig. 203).

En este caso, al aplicar la regla de reducción de la integral doble a una reiterada, el teorema de Newton — Leibniz (p. 29.3) y la fórmula (47.9), tenemos

$$\iint\limits_G \frac{\partial P}{\partial y} dxdy = \int\limits_a^b \left[\int\limits_{\varphi(x)}^{\psi(x)} \frac{\partial P}{\partial y} dy\right] dx = \int\limits_a^b \{P[x, \psi(x)] - P(x, \varphi(x))\} dx =$$

$$= \int\limits_a^b P[x, \psi(x)]dx - \int\limits_a^b P[x, \varphi(x)]dx = \int\limits_{\overset{\frown}{DC}} P(x, y)dx -$$

$$- \int\limits_{\overset{\frown}{AB}} P(x, y)dx = - \int\limits_{\overset{\frown}{CD}} P(x, y)dx - \int\limits_{\overset{\frown}{AB}} P(x, y)dx. \tag{47.13}$$

*) G. Green (1793 — 1841), matemático inglés.

**) Esto significa que $\{\overline{G}_i\}_{i=1}^{i=k}$ es la partición de la región cerrada $\overline{G}$ (véase el p. 44.3).

***) La continuidad de las derivadas parciales en $\overline{G}$ se entiende como su continuidad en el conjunto abierto G y su proiongabilidad continua a la frontera de G (véase el p. 39.3).

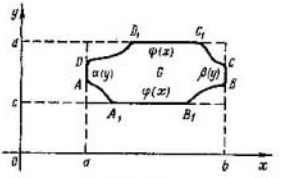


Fig. 203

Al notar que para los segmentos BC y DA

$$\int_{BC} P(x, y)dx = \int_{DA} P(x, y)dx = 0 \tag{47.14}$$

(esto proviene en seguida, por ejemplo, de la fórmula (47.6), pues aquí $x = \text{const}$ y, por ello, $\cos\alpha = 0$) y al sumar las igualdades (47.13) y (47.14), obtendremos

$$\iint_G \frac{\partial P}{\partial y}dxdy = -\int_{\overset{\frown}{AB}} Pdx - \int_{\overset{\frown}{BC}} Pdx - \int_{\overset{\frown}{CD}} Pdx -$$

$$\int_{\overset{\frown}{DA}} Pdx = -\int_{\gamma^+} Pdx. \tag{47.15}$$

Se ha obtenido tal orientación del contorno de frontera γ que los puntos A, B, C, D siguen uno tras otro sucesivamente. Esta orientación es positiva (véase la definición 6) y se designa mediante γ^+

De modo análogo, a partir de que la región G es elemental, se deduce la fórmula

$$\iint_G \frac{\partial Q}{\partial x}dxdy = \int_{\gamma^+} Qdy. \tag{47.16}$$

Al sumar (47.15) y (47.16), obtendremos la fórmula de Green (47.12) para el caso que se considera.

Examinemos el caso general. Supongamos que la región G está dividida en las regiones G_i $i = 1, 2, \ldots, k$, del tipo citado en las condiciones del teorema. En virtud de lo demostrado, para todo $i = 1, 2, \ldots, k$

$$\iint_{G_i} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dxdy = \int_{\gamma_i^+} Pdx + Qdy.$$

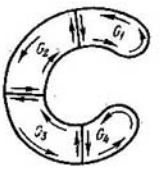


Fig. 204

Al sumar estas igualdades, obtendremos

$$\sum_{i=1}^{k} \iint_{G_i} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \sum_{i=1}^{k} \int_{\gamma_i^+} P dx + Q dy. \qquad (47.17)$$

Ya que la integral doble es aditiva respecto de los conjuntos (véase el p. 44.6), tendremos

$$\sum_{i=1}^{k} \iint_{G_i} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \iint_{G} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy. \qquad (47.18)$$

En la suma que figura en el segundo miembro de la igualdad (47.17) las integrales curvilíneas se toman dos veces a lo largo de todas las partes interiores de las fronteras γ_i de las regiones G_i, es decir, a lo largo de tales arcos de las curvas γ_i que forman parte de las fronteras de dos regiones G_i, $i = 1, 2, \ldots, k$, y, por consiguiente, no integran la frontera de la región G; además, las orientaciones de estos arcos de las curvas γ_i son opuestas (fig. 204). Puesto que una integral curvilínea de segunda especie cambia de signo cuando varía la orientación de la curva, la suma de dos integrales curvilíneas a lo largo de las partes indicadas de las curvas γ_i es nula. Por eso, en la suma derecha de la fórmula (47.17) quedarán sólo las integrales a lo largo de las partes orientadas positivas de la frontera γ de la región G, que dan en suma $\int_{\gamma^+} P dx + Q dy$. De este modo,

$$\sum_{i=1}^{k} \int_{\gamma_i^+} P dx + Q dy = \int_{\gamma^+} P dx + Q dy. \qquad (47.18')$$

De (47.17), (47.18) y (47.18') se deduce la fórmula (47.12) para el caso general. □

Sea G una región acotada en el plano R^2 y supongamos que su frontera consta de un número finito de contornos sencillos que se llamarán *contornos de frontera*. Si un contorno de frontera es, a la vez, la frontera de una región no acotada dispuesta en $R^2 \setminus \bar{G}$, se denominará *exterior*; en cambio, si es, a la vez, la frontera de

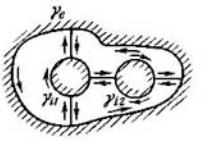


Fig. 205

una región acotada dispuesta en $R^2 \setminus \bar{G}$, se llamará *interior*. En la figura 205, γ_e es el contorno exterior, mientras que γ_{i1} y γ_{i2} son contornos interiores.

Si la frontera de la región G se compone del contorno exterior γ_e y los contornos interiores $\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{im}$ y si la región G puede ser dividida en un número finito de regiones, elementales respecto de ambos ejes de coordenadas, que tengan fronteras suaves a trozos, entonces resulta válida la fórmula

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy = \int_{\gamma_e^+} P\,dx + Q\,dy + \sum_{j=1}^{m} \int_{\gamma_{ij}^-} P\,dx + Q\,dy. \quad (47.19)$$

Como hasta ahora, las funciones P y Q se suponen continuas, junto con sus derivadas $\dfrac{\partial P}{\partial y}$ y $\dfrac{\partial Q}{\partial x}$, en la región cerrada $\bar{G}$.

Esta fórmula se demuestra igual que la (47.12), siempre que se toma en consideración que en la suma que figura en el segundo miembro de la igualdad (47.17) quedarán las integrales curvilíneas a lo largo de las partes orientadas positivas del contorno exterior y a lo largo de las partes orientadas negativas de los contornos interiores (véase la fig. 205).

Indiquemos, además, que en la fórmula (47.19) todos los contornos (tanto exteriores, como interiores) están orientados de una manera tal que durante su recorrido la región de integración queda a la izquierda.

Definición 7. *Supongamos que la frontera ∂G de una región plana acotada G se compone de un número finito de contornos sencillos suaves a trozos. La totalidad de estos contornos orientados de manera tal que, al recorrer cada uno de ellos, la región G queda a la izquierda (a la derecha), se llama orientación positiva (negativa) de la frontera de G y se designa $\partial G (-\partial G$, respectivamente).*

La fórmula de Green puede ser extendida a una clase de regiones más amplia. Con este fin diremos que en virtud de lo demostrado, la fórmula de Green es válida para un triángulo y, por ende, para cualquier polígono. Por eso, mediante un paso límite, aproximando la frontera de la región con unas quebradas compuestas de un número finito de lados, podemos obtener la fórmula de Green para cualquier región (e, incluso, para un conjunto abierto), cuya frontera se compone de un número finito de curvas suaves a trozos. Sin embargo, no nos detendremos en la demostración de este hecho y sólo nos limitaremos a su enunciación. Con ello, haciendo uso de la definición 7, escribiremos la fórmula (47.19) en una forma más compacta.

Teorema 1 . *Supongamos que la frontera de una región plana acotada G se compone de un número finito de curvas suaves a trozos. Si las funciones P, Q, $\frac{\partial P}{\partial y}$ y $\frac{\partial Q}{\partial x}$ son continuas en $\bar{G}$, entonces*

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \int_{\partial G} Pdx + Qdy,$$

donde ∂G es la frontera orientada positiva de la región G.

La fórmula de Green para las integrales múltiples sirve como análogo de la fórmula de Newton — Leibniz para integrales de multiplicidad unitaria: en ambas fórmulas las integrales de las derivadas extendidas a una región da integración se expresan a través de los valores de la función en la frontera de la región mencionada (en el caso de la fórmula de Green dichos valores se siguen integrando).

Ejercicios. Sirviéndose de la fórmula de Green, calcúlense las siguientes integrales curvilíneas, donde Γ es un contorno sencillo cerrado compuesto de las partes de las curvas cuyas ecuaciones están indicadas para cada integral (el sentido del recorrido del contorno es positivo).

5. $\int_\Gamma (x^2y + x - y)dx + (y^2 + 2x)dy;\ y = 2,\ y = x^2 + 1.$

6. $\int_\Gamma \frac{y}{x}dx + 2\ln x dy;\ 2x + y = 4,\ x = 1,\ y = 0.$

7. $\int_\Gamma \frac{dx}{y^2} - \frac{dy}{x};\ y = x,\ x = 2,\ y = 1.$

47.6. CÁLCULO DE LAS ÁREAS MEDIANTE INTEGRALES CURVILÍNEAS

Al poner en la fórmula de Green $Q = x$, $P = 0$, obtendremos $\iint_G dxdy = \int_{\gamma^+} xdy$, y, por lo tanto,

$$\mu G = \int_{\gamma^+} xdy. \tag{47.20}$$

Por analogía, al poner $P = -y$, $Q = 0$, obtendremos

$$\mu G = -\int_{\gamma^+} ydx. \tag{47.21}$$

Sumando las fórmulas (47.20) y (47.21), tendremos

$$\mu G = \frac{1}{2}\int_{\gamma^+} xdy - ydx. \tag{47.22}$$

Ejemplos. 1. Hallemos, con ayuda de la fórmula (47.22), el área limitada por la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Hagamos uso de su representación paramétrica: $x = a\cos t$, $y = b\operatorname{sen} t$. Al aplicar la fórmula (47.22), obtendremos el área buscada:

$$S = \frac{1}{2} \int_{\gamma^+} x\,dy - y\,dx = \frac{1}{2}\,ab \int_0^{2\pi} (\cos^2 t + \operatorname{sen}^2 t)\,dt = \pi ab.$$

Comparando este método de cálculo del área limitada por una elipse con el citado anteriormente (véase el ejemplo 4 en el p. 32.1), podemos cerciorarnos con facilidad cuán menos es aquí el volumen de cálculos.

2. Hallemos el área limitada por la astroide (véase en el tomo 1, fig. 84) $x = a\cos^3 t$, $y = a\operatorname{sen}^3 t$, $0 \leqslant t \leqslant 2\pi$. Teniendo en cuenta que en este caso el crecimiento del parámetro t corresponde a la orientación positiva del contorno, tenemos:

$$S = \frac{1}{2} \int_{\gamma^+} x\,dy - y\,dx = \frac{3a^2}{2} \int_0^{2\pi} (\cos^4 t \operatorname{sen}^2 t + \operatorname{sen}^4 t \cos^2 t)\,dt =$$

$$= \frac{3a^2}{8} \int_0^{2\pi} \operatorname{sen}^2 2t\,dt = \frac{3a^2}{16} \int_0^{2\pi} (1 - \cos 4t)\,dt = \frac{3\pi a^2}{8}$$

Ejercicios. Calcúlese, con ayuda de la integral curvilínea, el área de una figura limitada por las líneas:

8. $x = t^2$, $y = t^3$, $x = 1$, $y = 0$.
9. $y^2 = 4 - x$, $x = 1$, $y = 1$ (en el primer cuadrante).

47.7. SIGNIFICADO GEOMÉTRICO DEL SIGNO DEL JACOBIANO DE LA APLICACIÓN DE UNA REGIÓN PLANA

Sea F una aplicación biunívoca y continuamente derivable de la región plana $G \subset R^2_{uv}$ en el plano R^2_{xy} con un jacobiano distinto de cero en todo punto de G. En este caso, en virtud del principio de conservación de la región, el conjunto $G^* = F(G)$ es también una región (véase el p. 41.8), mientras que el jacobiano, siendo continuo, mantiene invariable su signo en G (véase el teorema 4 en el p. 19.6), es decir, o bien es G positivo en todo punto, o bien es negativo en todo punto.

Supongamos que en la inscripción de coordenadas la aplicación F se da mediante las fórmulas

$$x = x(u, v),$$
$$y = y(u, v). \tag{47.23}$$

Lema 1. *Si γ es una curva suave a trozos dispuesta en G, su imagen $\gamma^* = F(\gamma)$ será, realizándose la aplicación F, también una curva suave a trozos.*

DEMOSTRACIÓN. Supongamos al principio que γ es una curva suave, es decir, una curva continuamente derivable privada de puntos singulares (véase las definiciones 15 y 16 en el p. 16.4) y sea

$$u = u(t), \quad v = v(t), \quad a \leqslant t \leqslant b,$$

una representación suya. Entonces, en el segmento $[a, b]$ las funciones $u(t)$ y $v(t)$ son continuamente derivables y

$$\left(\frac{du}{dt}\right)^2 + \left(\frac{dv}{dt}\right)^2 > 0.$$

Como representación de la curva $\gamma^* = F(\gamma)$ servirá un par de funciones

$$x = x(u(t), v(t)), \quad y = y(u(t), v(t)), \quad a \leqslant t \leqslant b,$$

las cuales, en virtud de las propiedades de la composición de funciones continuamente derivables (véanse los pp. 19.3 y 20.3), serán también continuamente derivables. Mostremos que la curva γ^* no tendrá puntos singulares. Efectivamente, como

$$\frac{dx}{dt} = \frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt},$$

$$\frac{dy}{dt} = \frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt},$$

entonces, tomando estas igualdades por un sistema de ecuaciones lineales respecto de $\frac{du}{dt}$ y $\frac{dv}{dt}$, vemos que si en cierto punto $t \in [a, b]$ se cumplieran las igualdades $\frac{dx}{dt} = \frac{dy}{dt} = 0$, entonces, debido a que el jacobiano

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

no se reduce a cero, el sistema citado tendría la única solución, a saber, la solución nula, es decir, en el mismo punto t se verificarían las igualdades $\frac{du}{dt} = \frac{dv}{dt} = 0$, y, de este modo, el punto correspondiente de la curva γ sería singular lo que, por hipótesis, no es posible.

Así pues, si la curva γ es suave, la curva $\gamma^* = F(\gamma)$ es también suave. De aquí proviene en seguida que la imagen de una curva suave a trozos en la aplicación que se considera es también una curva suave a trozos, pues tal curva suave a trozos (véase la definición 16 en el p. 16.4) representa una reunión de un número finito de curvas suaves. □

Sea, ahora, $\Gamma \subset G$, donde Γ es una región limitada y su frontera $\partial\Gamma$ está constituida por el contorno sencillo suave a trozos (las fronteras de este tipo se llaman *suaves a trozos*). Supongamos, además, $\Gamma^* = F(\Gamma)$. Entonces, en virtud del principio de conservación de región, el conjunto Γ^* será también una región y, más aún, su frontera $\partial\Gamma^*$ es la imagen de la frontera $\partial\Gamma$ de la región Γ (véase el lema 1 en el p. 46.1), es decir, $\partial\Gamma^* = F(\partial\Gamma)$. Por eso la frontera $\partial\Gamma^*$ es también el contorno γ^* sencillo (debido al carácter biunívoco de la aplicación F) y suave a trozos (de acuerdo con el lema 1 de este punto). Por consiguiente, a lo largo de los contornos $\gamma = \partial\Gamma$ y $\gamma^* = \partial\Gamma^*$ pueden calcularse integrales curvilíneas. Supongamos que las regiones Γ y Γ^* son de tal índole que admiten el uso de la fórmula de Green, por ejemplo, satisfacen las condiciones que se imponen sobre la región en el teorema 1 del p. 47.5. (Según se ha señalado, la fórmula de Green, en realidad, es siempre aplicable bajo las suposiciones admitidas, pero esto no fue demostrado).

Designemos, como siempre, por γ^+ el contorno γ orientado positivo (véase el p. 47.5). Sea

$$u = u(t), \quad v = v(t), \quad a \leqslant t \leqslant b,$$

la representación del contorno γ^+ y, por lo tanto,

$$x = x[u(t), v(t)], \quad y = y[u(t), v(t)], \quad a \leqslant t \leqslant b, \tag{47.24}$$

es una representación del contorno γ^*.

Supondremos, además, que existen derivadas mixtas $\dfrac{\partial^2 y}{\partial v \partial u}$ y $\dfrac{\partial^2 y}{\partial u \partial v}$ y que éstas son continuas, por lo cual son iguales entre sí en todos los puntos de la región G.

De acuerdo con la fórmula (47.20),

$$\mu\Gamma^* = \varepsilon \int_{\gamma^*} x\,dy, \tag{47.25}$$

donde $\varepsilon = +1$, si la orientación del contorno γ^* es positiva y $\varepsilon = -1$, en el caso contrario. En otras palabras, $\varepsilon = +1$ ($\varepsilon = -1$, respectivamente), siempre que al recorrido positivo del contorno dado γ corresponde en la aplicación (47.23) el recorrido también positivo (negativo, respectivamente) del contorno $\gamma^* = F(\gamma)$.

Al transformar la integral (47.25) según la fórmula (47.8) y hacer uso de la representación (47.24) del contorno γ^*, tendremos

$$\mu\Gamma^* = \varepsilon \int_a^b xy'_t\,dt = \varepsilon \int_a^b x\left(\frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}\right)dt = \varepsilon \int_{\gamma^+} x\frac{\partial y}{\partial u}du + x\frac{\partial y}{\partial v}dv.$$

Apliquemos a la integral obtenida la fórmula de Green (véase el teorema 1 en el p. 47.5). Al poner $P = x\dfrac{\partial y}{\partial u}$, $Q = x\dfrac{\partial y}{\partial v}$ y observar que en este caso

$$\frac{\partial Q}{\partial u} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} + x\frac{\partial^2 y}{\partial u \partial v}, \quad \frac{\partial P}{\partial v} = \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} + x\frac{\partial^2 y}{\partial v \partial u}$$

(aquí se utiliza la condición de existencia de las segundas derivadas parciales $\dfrac{\partial^2 y}{\partial v \partial u}$

y $\frac{\partial^2 y}{\partial u \partial v}$, requerida anteriormente), obtendremos

$$\frac{\partial Q}{\partial u} - \frac{\partial P}{\partial v} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \frac{\partial (x, y)}{\partial (u, v)},$$

de donde

$$\mu\Gamma^* = \varepsilon \int_{\gamma^+} P du + Q dv = \varepsilon \iint_{\Gamma} \left(\frac{\partial Q}{\partial u} - \frac{\partial P}{\partial v}\right) du dv = \varepsilon \iint_{\Gamma} \frac{\partial (x, y)}{\partial (u, v)} du dv.$$

El primer miembro de esta igualdad es superior a cero, por lo tanto, el segundo miembro es también positivo y, como el jacobiano de la aplicación (47.23) no cambia de signo, esto es posible sólo en el caso en que $\varepsilon \frac{\partial (x, y)}{\partial (u, v)} > 0$, es decir, cuando el número ε es del mismo signo que el jacobiano $\frac{\partial (x, y)}{\partial (u, v)}$ y, en este caso, $\varepsilon \frac{\partial (x, y)}{\partial (u, v)} = \left|\frac{\partial (x, y)}{\partial (u, v)}\right|$. De este modo, el signo de ε no depende de la elección del contorno γ, sino que se determina por el signo del jacobiano el cual es el mismo en todos los puntos de la región G.

Queda demostrado, pues, el siguiente teorema.

Teorema 2. *Si las suposiciones asumidas anteriormente están cumplidas, resulta válida la fórmula*

$$\mu\Gamma^* = \iint_{\Gamma} \left|\frac{\partial (x, y)}{\partial (u, v)}\right| du dv. \tag{47.26}$$

Además, si $\frac{\partial (x, y)}{\partial (u, v)} > 0$ *en* Γ, *entonces* $\varepsilon = +1$; *en otras palabras, si el jacobiano de la aplicación F es positivo, entonces al recorrido positivo de todo contorno* $\gamma \subset G$, *que sirve de frontera para la región limitada* $\Gamma \subset G$, *le corresponde, en la aplicación F, el recorrido positivo del contorno* $\gamma^* = F(\gamma)$ *que sirve de frontera de la región limitada* $\Gamma^* = F(\Gamma)$. *En cambio, si el jacobiano* $\frac{\partial (x, y)}{\partial (u, v)} < 0$ *en* Γ, *entonces* $\varepsilon = -1$, *es decir, al recorrido positivo de todo contorno* γ *del tipo indicado le corresponde, en la aplicación F, un recorrido negativo del contorno* $\gamma^* = F(\gamma)$.

De este modo, *el significado geométrico del signo de un jacobiano consiste en que, siendo positivo el jacobiano, la orientación del contorno se mantiene invariable durante la aplicación y sí cambia, cuando el jacobiano es negativo.*

Con ayuda de la fórmula de Green (47.19) la fórmula (47.26) se generaliza fácilmente para el caso cuando la frontera de la región Γ se compone de un número finito de contornos cerrados suaves a trozos.

Diremos, además, que mediante la fórmula (47.26) puede obtenerse sin dificultad alguna la demostración más fácil del teorema 1, citado en el p. 46.1, sobre el significado geométrico del módulo del jacobiano. En efecto, sea $M_0 \in \Gamma$, $d(\Gamma)$ es el

diámetro de la región Γ, y supongamos que la región Γ se contrae de tal o cual modo hacia el punto M_0, y, por consiguiente, $d(\Gamma) \to 0$. Según el teorema del valor medio (véase el p. 44.6),

$$\mu\Gamma^* = \iint_{\Gamma} \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\,dv = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M \mu\Gamma, \quad M \in \Gamma,$$

por lo cual

$$\frac{\mu\Gamma^*}{\mu\Gamma} = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M.$$

Ya que el jacobiano es continuo, tenemos

$$\lim_{d(\Gamma)\to 0} \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_M = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_{M_0}$$

y, por lo tanto,

$$\lim_{d(\Gamma)\to 0} \frac{\mu\Gamma^*}{\mu\Gamma} = \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_{M_0}, \tag{47.27}$$

es decir, hemos demostrado la fórmula (46.9) y lo hicimos, en cierto sentido, de modo más general; así, por ejemplo, aquí Γ no es necesariamente un cuadrado (aunque, hemos impuesto sobre la aplicación F las condiciones un tanto más fuertes, al exigir que sean continuas las derivadas parciales $\frac{\partial^2 y}{\partial u \partial v}$ y $\frac{\partial^2 y}{\partial v \partial u}$, y, además, que se pueda aplicar la fórmula de Green para la región Γ^*). No es difícil convencerse de que la razón en el primer miembro de la fórmula (47.27) tiende a su límite uniformemente en el sentido indicado en el teorema 1 del p. 46.1.

Pese a que la deducción de la fórmula (47.27) resulta más fácil, lo que se ha conseguido a cuenta de las suposiciones más fuertes, cabe señalar que la demostración del teorema 1, aducida en el p. 46.1, es preferible desde el punto de vista del sentido matemático, puesto que revela mejor la misma esencia del problema vinculada con el hecho de que una aplicación derivable localmente se aproxima bastante bien por la aplicación lineal.

47.8. CONDICIONES DE INDEPENDENCIA DE UNA INTEGRAL CURVILÍNEA DEL CAMINO DE INTEGRACIÓN

Todas las curvas (los contornos) que se consideran en este punto se supondrán suaves a trozos lo que, para abreviar, no se especificará cada vez especialmente. Indiquemos, además, que en toda región G dos de sus puntos cualesquiera se pueden siempre unir con una curva suave a trozos, por ejemplo, con una quebrada (véase el lema 5 en el p. 41.4) dispuesta íntegramente en G.

Sea dada una región G y sean definidas en ésta las funciones continuas $P = P(x, y)$ y $Q = Q(x, y)$. Examinemos la cuestión concerniente a las condiciones que deben cumplirse para que la integral curvilínea $\int_{\overset{\frown}{AB}} P\,dx + Q\,dy$ (los puntos

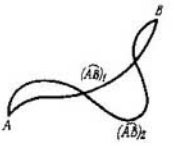


Fig. 206

$A \in G$ y $B \in G$ son arbitrariamente fijados) no dependa de la elección de la curva $\widehat{AB}$ que une A y B y que se dispone en G.

Lema 2. *La condición de independencia de la integral curvilínea en consideración del camino indicado de integración es equivalente a que sea nula la integral a lo largo de cualquier contorno cerrado dispuesto en la región G.*

DEMOSTRACIÓN. 1. En efecto, supongamos que para cualquier contorno cerrado $\gamma \subset G$ tiene lugar la igualdad

$$\int_{\gamma} Pdx + Qdy = 0$$

y sean dadas dos curvas $(\widehat{AB})_1$ y $(\widehat{AB})_2$ que unen en G los puntos A y B (véase la fig. 206). Designemos mediante $(\widehat{BA})_2$ una curva que se obtiene de $(\widehat{AB})_2$ al cambiar en ésta su orientación por la opuesta. La reunión $(\widehat{AB})_1 \cup (\widehat{BA})_2$ de las curvas $(\widehat{AB})_1$ y $(\widehat{BA})_2$ es un contorno cerrado, razón por la cual

$$\int_{(\widehat{AB})_1 \cup (\widehat{BA})_2} Pdx + Qdy = 0; \tag{47.28}$$

pero

$$\int_{(\widehat{AB})_1 \cup (\widehat{BA})_2} Pdx + Qdy = \int_{(\widehat{AB})_1} Pdx + Qdy +$$

$$+ \int_{(\widehat{BA})_1} Pdx + Qdy = \int_{(\widehat{AB})_1} Pdx + Qdy - \int_{(\widehat{AB})_2} Pdx + Qdy. \tag{47.29}$$

De (47.28) y (47.29) proviene que

$$\int_{(\widehat{AB})_1} Pdx + Qdy = \int_{(\widehat{AB})_2} Pdx + Qdy,$$

es decir, la integral curvilínea $\int_{\widehat{AB}} Pdx + Qdy$ no depende del camino de integración $\widehat{AB} \subset G$, cuando $A \in G$ y $B \in G$ están fijados.

2. Viceversa, supongamos que la integral $\int Pdx + Qdy$ no depende del camino de integración en el sentido mencionado y sea dado un contorno cerrado γ dis-

puesto en G. Elijamos en éste dos puntos A y $B \neq A$; entonces $\gamma = \widehat{AB} \cup \widehat{BA}$ y

$$\int_{\gamma} Pdx + Qdy = \int_{\widehat{AB}} + \int_{\widehat{BA}} = \int_{\widehat{AB}} - \int_{(\widehat{AB})_1} = 0,$$

donde $(\widehat{AB})_1$ denota la curva que se obtiene de la curva $\widehat{BA}$, al cambiar en ésta su orientación por la opuesta. □

Enunciemos el criterio de independencia de una integral del camino de integración.

Teorema 3. *Supongamos que las funciones $P(x, y)$ y $Q(x, y)$ son continuas en la región plana G. Para que la integral curvilínea $\int_{\widehat{AB}} Pdx + Qdy$, con los puntos fijados $A \in G$ y $B \in G$, no dependa del camino de integración $AB \subset G$, es necesario y suficiente que la expresión $Pdx + Qdy$ sea la diferencial total de cierta función $u = u(x, y)$, definida en la región G:*

$$du = Pdx + Qdy \tag{47.30}$$

(esto es equivalente a que $\frac{\partial u}{\partial x} = P, \frac{\partial u}{\partial y} = Q, (x, y) \in G$).

Si esta condición se cumple para cualesquiera dos puntos $A = (x_0, y_0) \in G$ y $B = (x_1, y_1) \in G$ y toda curva $\widehat{AB}$ que une dichos puntos en G: $\widehat{AB} \subset G$, se verifica la identidad

$$\int_{\widehat{AB}} Pdx + Qdy = u(x_1, y_1) - u(x_0, y_0). \tag{47.31}$$

DEMOSTRACIÓN DE LA NECESIDAD DE LA CONDICIÓN (47.30). Admitamos que la integral en consideración no depende del camino de integración dispuesto en la región G, sino que sólo depende de los puntos inicial y final del camino. Supongamos que $M_0 = (x_0, y_0) \in G$, $M = (x, y) \in G$ y $\widehat{M_0M}$ es una curva suave a trozos que une en G los puntos M_0 y M (una curva de tal género, incluso una quebrada, siempre existe, véase el lema 5 del p. 41.4). Pongamos

$$u(M) = u(x, y) \overset{\text{def}}{=} \int_{\widehat{M_0M}} Pdx + Qdy.$$

La función $u(x, y)$ es unívoca, puesto que el valor $u(M) = u(x, y)$ no depende de la elección de la curva que une en G los puntos M_0 y M. Probemos que

$$\frac{\partial u(x, y)}{\partial x} = P(x, y) \quad \text{y} \quad \frac{\partial u(x, y)}{\partial y} = Q(x, y).$$

Fijemos el punto $M = (x, y)$ y elijamos el punto $M_h = (x + h, y) \in G, h \neq 0$, de una manera tal que el segmento MM_h que une M y M_h (el cual, evidentemente,

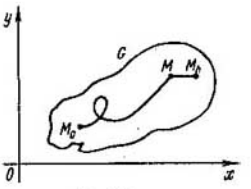


Fig. 207

es paralelo al eje Ox y tiene la longitud igual a $|h|$), se contenga en G (fig. 207). Para todos los números h, suficientemente pequeños, tal elección es siempre posible (¿por qué?). En este caso tenemos

$$u = (x + h, y) - u(x, y) =$$

$$= \int_{\widehat{M_0M_h}} P dx + Q dy - \int_{\widehat{M_0M}} P dx + Q dy = \int_{\widehat{MM_h}} P dx + Q dy.$$

A lo largo del segmento MM_h la coordenada y se mantiene constante, por lo cual $\int_{MM_h} Q dy = 0$, y, por consiguiente, $u(x + h, y) - u(x, y) = \int_{MM_h} P dx =$ $= \int_x^{x+h} P(t, y) dt$. Al aplicar el teorema integral del valor medio, obtendremos

$$u(x + h, y) - u(x, y) = P(x + \theta h, y)h, \quad 0 < \theta < 1,$$

de donde

$$\frac{u(x + h, y) - u(x, y)}{h} = P(x + \theta h, y), \quad 0 < \theta < 1. \tag{47.32}$$

El segundo miembro de esta igualdad tiene un límite para $h \to 0$, puesto que la función $P(x, y)$ es continua, y por consiguiente el primer miembro también tiene un límite cuando $h \to 0$. Pasando al límite en (47.32), tendremos $\frac{\partial u(x, y)}{\partial x} = P(x, y)$.

De un modo completamente análogo se demuestra también la igualdad $\frac{\partial u(x, y)}{\partial y} = Q(x, y)$.

Así pues, la existencia de la función $u(x, y)$, para la cual se verifica la correlación (47.30), queda demostrada.

Supongamos ahora que $A \in G$, $B \in G$, $\widehat{AB}$ es una curva que une en G los puntos A y B, y sea $x = x(t), y = y(t), a \leqslant t \leqslant b$, una representación suya y, por lo tan-

to, $A = (x(a), y(a))$, $B = (x(b), y(b))$. En este caso

$$\int_{\overset{\frown}{AB}} P dx + Q dy = \int_a^b [P[x(t), y(t)]x'(t) + Q[x(t), y(t)]y'(t)]dt =$$

$$= \int_a^b \left[\frac{\partial u(x(t), y(t))}{\partial x}\frac{dx}{dt} + \frac{\partial u(x(t), y(t))dy}{\partial y}\frac{dy}{dt}\right]dt =$$

$$= \int_a^b u_t(x(t), y(t))dt = u[x(b), y(b)] - u[x(a), y(a)] = u(B) - u(A),$$

es decir, la fórmula (47.31) queda también demostrada.

LA DEMOSTRACIÓN DE SUFICIENCIA de la condición (47.30) referente a la independencia de una integral curvilínea del camino de integración se deduce directamente de la fórmula (47.31). Efectivamente, el punto inicial de cualquier contorno cerrado γ coincide con el final, por lo cual, en virtud de (47.31), se tiene

$$\int_\gamma P dx + Q dy = u(A) - u(A) = 0.$$

De conformidad con el lema 2, esto precisamente significa que la integral curvilínea correspondiente no depende del trayecto de integración. □

Observemos que, aunque el teorema demostrado nos proporciona las condiciones necesarias y suficientes referentes a la independencia de una integral curvilínea del camino de integración, dichas condiciones se comprueban con dificultad.

Al hacer más estrecha la clase de las regiones en consideración podemos obtener un criterio considerablemente más simple y efectivo. Introduzcamos la siguiente definición.

Definición 8. *Una región plana G se llama simplemente conexa, si, cualquiera que sea el contorno sencillo $\gamma \subset G$, la región limitada Γ, cuya frontera está constituida por γ, se contiene en G.*

Hablando metafóricamente, el hecho de que una región es simplemente conexa significa que ella no tiene "agujeros". Un círculo nos da el ejemplo de una región simplemente conexa, un anillo circular representa una región no simplemente conexa (fig. 208).

Antes de enunciar el otro criterio que caracteriza la independencia de la integral curvilínea del camino de integración, demostremos un lema que nos hará falta al demostrar dicho criterio.

Lema 3. *Supongamos que las funciones $P(x, y)$ y $Q(x, y)$ son continuas en la región G; γ es una curva suave dispuesta en G; $x = x(t), y = y(t), a \leqslant t \leqslant b$, es su representación, $\tau = \{t_i\}_{i=0}^{i=i_0}$ es la partición del segmento $[a, b]$, y λ_τ, una quebrada cuyos vértices se ubican en los puntos $(x(t_i), y(t_i))$, $i = 0, 1, \ldots, i_0$* (véase el p. 16.5). *En este caso*

$$\lim_{\delta_\tau \to 0} \int_{\lambda_\tau} P dx + Q dy = \int_\gamma P dx + Q dy. \qquad (47.33)$$

Fig. 208

Cabe señalar que como las funciones $x(t)$ e $y(t)$ son uniformemente continuas en el segmento $[a, b]$, las longitudes de los lados de la quebrada λ_τ, es decir, las de los segmentos cuyos vértices se encuentran en los puntos $(x(t_{i-1}), y(t_{i-1}))$ y $(x(t_i), y(t_i))$ tienden también a cero cuando $\delta_\tau \to 0$.

DEMOSTRACIÓN. La curva γ es un compacto; ya que este compacto no se interseca con el conjunto cerrado $R^2_{xy} \setminus G$, la distancia entre ellos es superior a cero (véase el lema 7 del p. 18.2). Sea η un número cualquiera tal que $\rho(\gamma, R^2_{xy} \setminus G) > \eta > 0$. Designemos mediante γ_η la totalidad de todos los puntos del plano que distan de γ a una magnitud no superior a η. El conjunto γ_η está acotado, cerrado (véase el p. 18.3, lema 11) y $\gamma_\eta \subset G$.

Dado que las funciones $x(t)$ e $y(t)$ son uniformemente continuas en el segmento $[a, b]$, existe tal número $\delta_0 > 0$, que para cualesquiera dos puntos $t' \in [a, b]$ y $t'' \in [a, b]$, que satisfacen la condición $|t'' - t'| < \delta_0$, se cumple la desigualdad

$$\rho(M', M'') = \sqrt{[x(t'') - x(t')]^2 + [y(t'') - y(t')]^2} < \eta,$$

donde $M' = (x(t'), y(t'))$, $M'' = (x(t''), y(t''))$ (compárese con el lema 4 en el p. 41.4). Es evidente que todos los puntos del segmento cuyos extremos se ubican en los puntos M' y M'' se hallan alejados del punto M' a una distancia no superior a η, razón por la cual se disponen en el conjunto γ_η, y, por consiguiente, en G. Esto significa que si la finura δ_τ de la partición τ del segmento $[a, b]$ es tal que $\delta_\tau < \delta_0$, todos los puntos de la quebrada λ_τ se disponen en G y para las particiones de este género τ tiene sentido la integral $\int\limits_{\lambda_\tau} P dx + Q dy$.

Examinemos las integrales $\int\limits_\gamma P dx$ y $\int\limits_{\lambda_\tau} P dx$. Pongamos

$$x_i = x(t_i), \quad y_i = y(t_i), \quad P_i = P(x_i, y_i),$$

$$\Delta x_i = x_i - x_{i-1}, \; i = 1, 2, \ldots, i_0, \; \sigma_\tau = \sum_{i=1}^{i_0} P_i \Delta x_i.$$

Según se sabe (véase el p. 47.2, la propiedad 4°),

$$\lim_{\delta_\tau \to 0} \sigma_\tau = \int\limits_\gamma P dx. \tag{47.34}$$

Sean, luego, $M_i = (x_i, y_i)$ los vértices de la quebrada λ_τ; en este caso

$$\int\limits_{\lambda_\tau} P dx = \sum_{i=1}^{i_0} \int\limits_{M_{i-1}M_i} P dx. \tag{47.35}$$

Por otra parte, indiquemos que (haciendo uso de las designaciones adoptadas en el p. 47.2)

$$\int_{M_{i-1}M_i} dx = \int_{M_{i-1}M_i} \cos\alpha\, ds = |M_{i-1}M_i| \cos\alpha = \Delta x_i,$$

por lo cual

$$\sigma_\tau = \sum_i P_i \Delta x_i = \sum_i P_i \int_{M_{i-1}M_i} dx = \sum_i \int_{M_{i-1}M_i} P_i\, dx. \tag{47.36}$$

Al designar con L_τ la longitud de la quebrada λ_τ; con S, la longitud de la curva γ y con $\omega(\delta; P)$, el módulo de continuidad de la función $P(x, y)$ en el compacto γ_η, y teniendo presente que en virtud de la definición de la curva: $L_\tau \leqslant S$, obtendremos, a partir de (47.35) y (47.36):

$$\left| \int_{\lambda_\tau} P\,dx - \sigma_\tau \right| \leqslant \sum \left| \int_{M_{i-1}M_i} |P - P_i|\, dx \right| \leqslant$$

$$\leqslant \omega(\delta_\tau; P) \sum_i |\Delta x_i| \leqslant \omega(\delta_\tau; P) L_\tau \leqslant \omega(\delta_\tau; P) S.$$

De aquí, en virtud de que la función $P(x, y)$ es uniformemente continua en el conjunto γ_η, tenemos $\lim_{\delta_\tau \to 0} \left(\int_{\lambda_\tau} P\,dx - \sigma_\tau \right) = 0$, y, por ende, de conformidad con (47.34),

$$\lim_{\delta_\tau \to 0} \int_{\lambda_\tau} P\,dx = \int_\gamma P\,dx. \tag{47.37}$$

Análogamente se demuestra también la igualdad

$$\lim_{\delta_\tau \to 0} \int_{\lambda_\tau} Q\,dy = \int_\gamma Q\,dy. \tag{47.38}$$

De (47.37) y (47.38) se deduce directamente la afirmación del lema, es decir, la fórmula (47.33). □

OBSERVACIÓN. Una afirmación, análoga al lema, es lícita también para las integrales curvilíneas en un espacio, con la particularidad de que la demostración del caso espacial se realiza siguiendo el mismo esquema que se usa en el caso plano.

Teorema 4. *Supongamos que las funciones $P(x, y)$ y $Q(x, y)$ son continuas junto con sus derivadas parciales $\frac{\partial P}{\partial y}$ y $\frac{\partial Q}{\partial x}$ en una región plana G. Para que la integral curvilínea $\int_{\overset{\frown}{AB}} P\,dx + Q\,dy$ no dependa del camino de integración $\overset{\frown}{AB} \subset G$, siendo fijados arbitrariamente los puntos $A \in G$ y $B \in G$, es necesario y suficiente (siempre que la región G es simplemente conexa) que en todos los puntos de las región G se cumpla la igualdad $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.*

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que la integral en consideración no depende del camino de integración dispuesto en la región G, sino sólo de la posición de los puntos inicial y final. De acuerdo con el teorema 3, existe una función $u=u(x, y)$ tal que $du=Pdx+Qdy$, es decir, tal que $\frac{\partial u}{\partial x}=P$, $\frac{\partial u}{\partial y}=Q$. Como $\frac{\partial P}{\partial y}=\frac{\partial^2 u}{\partial y \partial x}$, $\frac{\partial Q}{\partial x}=\frac{\partial^2 u}{\partial x \partial y}$ y, por hipótesis del teorema, las derivadas $\frac{\partial P}{\partial y}$ y $\frac{\partial Q}{\partial x}$ y, por lo tanto, las derivadas mixtas $\frac{\partial^2 u}{\partial y \partial x}$ y $\frac{\partial^2 u}{\partial x \partial y}$ son continuas, entonces (véase el p. 21.1), son, además, iguales, es decir, $\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Supongamos ahora que la región G es simplemente conexa y en todos los puntos suyos se verifica $\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$. Si γ es un contorno cerrado sencillo y suave a trozos, dispuesto en G, y si D es una región limitada cuya frontera es γ, entonces, al aplicar la fórmula de Green (utilizamos aquí el hecho de que la región G es simplemente conexa), obtendremos

$$\int\limits_{\gamma^+} P dx + Q dy = \iint\limits_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx dy = 0.$$

Si la curva γ, dispuesta en G, tiene un número finito de puntos múltiples, entonces para cada uno de sus "bucles" γ_k, $k = 1, 2, \ldots, k_0$, que representa ya un contorno cerrado sencillo, tenemos sucesivamente, en virtud de lo demostrado, $\int\limits_{\gamma_k} P dx + Q dy = 0$, de donde se infiere que para toda la curva γ también

$$\int\limits_{\gamma} P dx + Q dy = 0. \tag{47.39}$$

Pasando al caso general, fijemos nuestra atención en que mediante el procedimiento considerado la igualdad (47.39) se demuestra también en el caso cuando γ sea una quebrada cerrada de un número finito de lados. Desde el punto de vista geométrico, la diferencia sólo consiste en que la autointersección de una quebrada de un número finito de lados puede componerse no sólo del número finito de puntos, sino también del número finito de segmentos lo que complica los razonamientos de un modo insignificante. La posibilidad de aplicar la fórmula de Green a una región finita, limitada por una quebrada del número finito de lados, proviene de que la región citada puede ser dividida en triángulos, los cuales son, evidentemente, regiones elementales respecto a ambos ejes de coordenadas. Por consiguiente, en este caso se cumplen las premisas del teorema 1 en el p. 47.3.

Entre tanto, cualquier curva cerrada suave a trozos γ, dispuesta en G, puede ser aproximada, con cualquier grado deseado de precisión, mediante unas quebradas cerradas de un número finito de lados, a consecuencia de lo cual, pasando al límite, podemos obtener la igualdad (47.39) para toda curva cerrada de G. Hagámoslo.

Supongamos que γ es una curva cerrada suave a trozos en la región G; la citada curva está dada mediante la representación $r(t)$, $a \leqslant t \leqslant b$, y es una reunión de las

curvas suaves $\gamma_1, \ldots, \gamma_k$. Inscribamos la quebrada λ_j dentro de toda curva γ_j, $j = 1, 2, \ldots, k$. La reunión de todas las quebradas λ_j, $j = 1, 2, \ldots, k$, forma una línea quebrada cerrada λ, correspondiente a cierta partición τ del segmento $[a, b]$. En virtud de lo demostrado, tenemos

$$\int_{\lambda} Pdx + Qdy = 0.$$

Pero, de acuerdo con el lema,

$$\lim_{\delta_\tau \to 0} \int_{\lambda_j} Pdx + Qdy = \int_{\gamma_j} Pdx + Qdy, \quad j = 1, 2, \ldots, k,$$

y, por lo tanto,

$$\lim_{\delta_\tau \to 0} \int_{\lambda} Pdx + Qdy = \int_{\gamma} Pdx + Qdy,$$

por lo cual

$$\int_{\gamma} Pdx + Qdy = 0. \square$$

La condición $\dfrac{\partial P}{\partial y} \dfrac{\partial Q}{\partial x}$ se denomina, a veces, *criterio de la diferencial total en una región simplemente conexa*, puesto que, de acuerdo con los teoremas 3 y 4, dicha condición es necesaria y suficiente para que la expresión $Pdx + Qdy$ en la región G sea una diferencial de cierta función $u(x, y)$, $(x, y) \in G$.

Como conclusión de este punto hemos de notar que al demostrar la suficiencia de las condiciones del teorema 4 para que la integral curvilínea sea independiente del camino de integración, resulta muy esencial la exigencia de que la región en consideración sea simplemente conexa y esta exigencia no puede ser menospreciada. Confirmemos esto con un ejemplo.

Ejemplo. Sea $P(x, y) = -\dfrac{y}{x^2 + y^2}$, $Q(x, y) = \dfrac{x}{x^2 + y^2}$.

Es fácil compobar que

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y} \tag{47.40}$$

para todos los puntos del plano, excepto el origen de coordenadas (0, 0). Esto proviene, por ejemplo, de que

$$d\left(\operatorname{arctg}\frac{y}{x}\right) = \frac{-ydx + xdy}{x^2 + y^2}, \quad x^2 + y^2 > 0. \tag{47.41}$$

De este modo, en el caso dado por la región G puede tomarse todo el plano con el origen de coordenadas "punzado": $G = R^2 \setminus \{(0, 0)\}$. Es evidente que la región G no es simplemente conexa. A título de contorno cerrado elijamos la circunferencia unidad

$$\gamma_0 = \{x = \cos t, y = \operatorname{sen} t, 0 \leqslant t \leqslant 2\pi\},$$

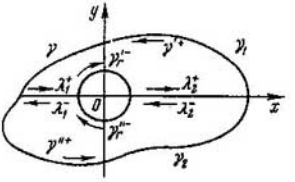


Fig. 209

entonces

$$\int_{\gamma_0} P\,dx + Q\,dy = \int_{\gamma_0} \frac{-y\,dx + x\,dy}{x^2 + y^2} = \int_0^{2\pi} dt = 2\pi.$$

Por consiguiente, en este caso las condiciones (47.40) quedan cumplidas y existe un contorno cerrado γ_0, a lo largo del cual la integral no es nula. No es difícil convencerse de que, en general,

$$\int_{\gamma_r} P\,dx + Q\,dy = 2\pi \tag{47.42}$$

a lo largo de cualquier circunferencia γ_r de radio r y con centro en el origen de coordenadas.

Luego, cualquiera que sea el contorno sencillo suave a trozos γ que constituya la frontera de una región limitada en la que está contenido el origen de coordenadas (en este caso suele decirse que el contorno γ contiene en su interior el origen de coordenadas), para este contorno se verifica también

$$\int_{\gamma} P\,dx + Q\,dy = 2\pi. \tag{47.43}$$

Para demostrarlo elijamos una circunferencia γ_r de tal radio r que sea $\gamma_r \subset \Gamma$; en este caso γ y γ_r no se intersecan. Al unir los contornos γ y γ_r por medio de los segmentos λ_1 y λ_2, como lo muestra fig. 209, obtendremos dos contornos cerrados γ_1 y γ_2 que no contienen en su interior el origen de coordenadas y que se componen de los arcos γ_r' y γ_r'' de la circunferencia γ_r, las partes γ' y γ'' del contorno γ y los segmentos λ_1 y λ_2.

En virtud de la condición (47.40), para estos contornos se verifican las igualdades

$$\int_{\gamma_1} P\,dx + Q\,dy = 0, \qquad \int_{\gamma_2} P\,dx + Q\,dy = 0.$$

Al sumar estas igualdades y omitir, para abreviar, las expresiones subintegrales, obtendremos (fig. 209):

$$0 = \int_{\gamma_1} + \int_{\gamma_2} = \int_{\gamma'} + \int_{\lambda_1^+} + \int_{\gamma_r'^-} + \int_{\lambda_2^+} + \int_{\gamma''} + \int_{\lambda_2^-} + \int_{\gamma_r''^-} + \int_{\lambda_1^-} =$$

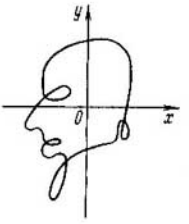


Fig. 210

$$= \int_{\gamma'+} + \int_{\gamma''+} - \int_{\gamma_r'+} - \int_{\gamma_r''+} = \int_{\gamma^+} - \int_{\gamma_r^+}$$

De aquí, debido a (47.42), proviene precisamente (47.43). Más aún, esta igualdad se verifica también en el caso, cuando el contorno γ, al recorrer "una vez" alrededor del origen de coordenadas, forma un número finito de "bucles" que no cercan el origen de coordenadas (fig. 210), pues la integral a lo largo de estos bucles es nula.

Si M_0 es un punto fijo de la región en consideración G, $M_0 \in G$, $M \in G$, y $\widehat{M_0M}$ es una curva que une en G los puntos M_0 y M, entonces $u(M) = \int\limits_{\widehat{M_0M}} P\,dx + Q\,dy$ ya será una función multiforme cuyos valores se determinan por la elección de distintos trayectos que unen los puntos M_0 y M. Si γ_0 es una curva fija cualquiera que une M_0 y M, entonces todos los valores de la función u en el punto M se dan mediante la fórmula

$$u(M) = \int_{\gamma_0} P\,dx + Q\,dy + 2\pi n, \quad n = 0, \pm 1, \pm 2, \ldots$$

esto es, cada recorrido en torno al origen de coordenadas cambia el valor de la función $u(M)$ en la magnitud $\pm 2\pi$, según sea el sentido del recorrido.

En el caso dado podemos convencernos de esto inmediatamente: de la fórmula (47.41) se deduce que

$$\int_{\gamma_0} P\,dx + Q\,dy = \int_{\gamma_0} \frac{-y\,dx + x\,dy}{x^2 + y^2} = \left(\operatorname{Arctg}\frac{y}{x}\right)_0,$$

donde $\left(\operatorname{Arctg}\dfrac{y}{x}\right)_0$ es un cierto valor fijo de $\operatorname{Arctg}\dfrac{y}{x}$; por esto

$$u(M) = \operatorname{Arctg}\frac{y}{x}.$$

Un lector reflexivo se dará cuenta de que muchos razonamientos aducidos en este ejemplo no dependen de la forma concreta de las funciones P y Q y son verídicos siempre cuando nos enfrentemos con un "punto singular" aislado, es decir, un punto en el que se perturba la condición (47.40). Por supuesto, "recorrido" tal punto singular una sóla vez, se obtendrá no 2π, sino, en el caso general, algún otro número.

singular una sóla vez, se obtendrá no 2π, sino, en el caso general, algún otro número.

El resultado análogo al teorema 4 tiene lugar también cuando γ es una curva espacial (véase el p. 52.6).

Ejercicios. 10. Demuéstrese la fórmula

$$\iint_G v\Delta u\,dx\,dy = -\iint_G \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dx\,dy + \int_{\gamma^+} v\frac{\partial u}{\partial \nu}\,ds,$$

donde G es una región plana, para la cual es válida la fórmula de Green, γ es el contorno que la limita, ν es la normal exterior unidad al contorno γ, y Δ, el operador de Laplace (véase el p. 41.10).

11. Calcúlese la integral $\int_\Gamma 2(x + y^2)dx + (4xy + \cos y)dy$, donde Γ es una curva arbitraria suave a trozos que une los puntos $(1, 0)$ y (ξ, η).

12. Sea Γ un contorno cerrado sencillo suave a trozos, arbitrariamente elegido, que limita una región en la que está contenido el origen de coordenadas. Calcúlese la integral

$$\int_\Gamma \frac{e^x}{x^2 + y^2}\left[(x\operatorname{sen} y - y\cos y)dx + (x\cos y + y\operatorname{sen} y)dy\right]$$

para el sentido positivo del recorrido del contorno Γ

§ 48. INTEGRALES MÚLTIPLES IMPROPIAS

48.1. DEFINICIONES FUNDAMENTALES

Al igual que en el caso de integrales ordinarias, introduzcamos el concepto de integral múltiple impropia, es decir, de integral múltiple de las funciones que o bien no están acotadas o bien están definidas en una región no limitada. La definición de la integral múltiple impropia la enunciemos de modo tal que se abarquen ambos casos mencionados (compárese con el p. 33.1).

Definición 1. *Sea G un conjunto abierto (acotado o no acotado) en el espacio n-dimensional R^n. Una sucesión de conjuntos abiertos G_k, $k = 1, 2, \ldots$, se llamará sucesión que agota de modo monótono el conjunto abierto G, si*

1) $\bar{G}_k \subset G_{k+1}$, $k = 1, 2, \ldots$;

2) $\bigcup_{k=1}^{\infty} G_k = G$.

Aquí $\bar{G}$ significa, como siempre, la clausura (véase el p. 18.2) del conjunto G.

Definición 2. *Sea dada en un conjunto abierto G la función f (acotada o no acotada) integrable según Riemann en cualquier conjunto abierto D, medible según Jordan, de tal género que $\bar{D} \subset G$. La función f se denomina integrable en el sentido impropio en el conjunto abierto G, siempre que para toda sucesión de conjuntos*

medibles abiertos G_k, $k = 1, 2, \ldots$, *que agota de modo monótono el conjunto* G, *existe el límite* $\lim_{k\to\infty} \int f dG_k$ *que no depende de la elección de la sucesión citada* G_k, $k = 1, 2, \ldots$.

Este límite recibe el nombre de integral impropia de la función f *extendida al conjunto abierto* G *y se designa por* $\int f dG$, *o, en la forma desarrollada*

$$\iint\limits_G \cdots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n.$$

De este modo,

$$\int f dG = \lim_{k\to\infty} \int f dG_k. \tag{48.1}$$

Si la integral $\int f dG$ *existe, suele decirse que ella converge, de lo contrario, la integral diverge.*

Cabe notar que en el caso de $n = 1$, la definición citada de la integral impropia no es equivalente a la definición de integral impropia de la función de una sola variable, dada en el § 33. Esto se debe a que en el párrafo indicado a título de conjuntos G_k se han tomado sólo los intervalos, es decir, conjuntos abiertos medibles unidimensionales del tipo muy especial. Por esta razón, el concepto de integral impropia (48.1), introducido en este párrafo, lo emplearemos sólo en el caso de $n \geqslant 2$, conservando intacto para el caso de $n = 1$ la noción anterior de la integral impropia.

Si el conjunto abierto G es medible según Jordan y la función f es integrable en G, la integral impropia de la función f coincide con la integral habitual de Riemann, lo que resulta de la aditividad completa de la integral de Riemann (véase el p. 44.6).

La definición (48.1) permite extender a las integrales impropias toda una serie de las propiedades de integrales propias: aditividad de la integral en los conjuntos, linealidad de la integral, integración de las desigualdades, reducción de la integral múltiple a una reiterada, fórmula de cambio de variable, etc.

Por ejemplo, si $x = F(u)$ es una aplicación biunívoca continuamente derivable del conjunto abierto $D \subset R^n_u$ sobre otro conjunto abierto $G \subset R^n_x$, y si el jacobiano $J(u)$ de dicha aplicación nunca se reduce a cero en D, entonces para toda función f, continua en G, queda válida la fórmula del cambio de variable en la integral:

$$\int f(x) dG = \int f[F(u)] |J(u)| dD.$$

La demostración de esta afirmación es la misma que la usada para el teorema 2 en el p. 46.2; conviene sólo usar la definición (48.1) en lugar de la aditividad completa de la integral.

Haciendo uso de la aditividad de la integral múltiple impropia, podemos escribir la definición (48.1) en otra forma equivalente. Teniendo presente que para un conjunto abierto medible $\Gamma \subset G$ es válida la igualdad

$$\int f dG - \int f d\Gamma = \int f d(G \setminus \overline{\Gamma}), \tag{48.2}$$

podemos decir que la integral $\int f dG$ converge cuando, y sólo cuando, para cual-

quier sucesión de conjuntos abiertos medibles G_k, $k = 1, 2, \ldots$, que agota de modo monótono el conjunto G, existan las integrales $\int f d(G \setminus \bar{G}_k)$ y

$$\lim_{k \to \infty} \int f d(G \setminus \bar{G}_k) = 0.$$

Ejercicio 1. Demuéstrese la fórmula (48.2); en particular, pruébese que las integrales $\int f dG$ y $\int f d(G \setminus \Gamma)$ convergen o divergen simultáneamente.

48.2. INTEGRALES IMPROPIAS DE LAS FUNCIONES NO NEGATIVAS

Teorema 1. *Sea f una función no negativa en el conjunto abierto $G \subset R^n$. Cualquiera que sea la sucesión $\{G_k\}$ de conjuntos abiertos medibles según Jordan G_k, que agota de modo monótono el conjunto G, el límite*

$$\lim_{k \to \infty} \int f(x) dG, \tag{48.3}$$

sea finito o igual a $+\infty$, siempre existe.

Si dicho límite es finito, *la integral* $\int f(x) dG$ *existe y, por lo tanto, el límite* (48.3) *es igual a esta integral; si, en cambio, el límite* (48.3) *es infinito, la integral* $\int f(x) dG$ *no existe.*

En este último caso se escribe $\int f(x) dG = +\infty$. Esto se justifica por lo que, en virtud del teorema enunciado, para cualquier otra sucesión $\{D_k\}$ de conjuntos abiertos medibles D_k, que agota de modo monótono G, tenemos

$$\lim_{k \to \infty} \int f(x) dD_k = +\infty.$$

DEMOSTRACIÓN. Es evidente que el teorema quedará demostrado, si probamos que bajo el supuesto de que en el conjunto abierto G la función f es no negativa, para cualquier sucesión de conjuntos medibles G_k, $k = 1, 2, \ldots$, que agota de modo monótono la región G, existe un límite finito o infinito

$$\lim_{k \to \infty} \int f dG_k$$

y este límite no depende de la elección de la sucesión citada.

Sea G_k, $k = 1, 2, \ldots$, una sucesión de conjuntos medibles que agota de modo monótono el conjunto abierto G. Entonces conforme a la definición de tal sucesión, $G_k \subset G_{k+1}$, y, como $f \geqslant 0$, se tiene $\int f dG_k \leqslant \int f dG_{k+1}$, $k = 1, 2, \ldots$, y, por ende, siempre existe un límite finito o infinito

$$\lim_{k \to \infty} \int f dG_k = I_1.$$

Sea ahora D_k, $k = 1, 2, \ldots$, alguna otra sucesión de conjuntos medibles que agota de modo monótono el conjunto abierto G. En virtud de lo demostrado más arriba, existe un límite finito o infinito

$$\lim_{k \to \infty} \int f dD_k = I_2.$$

Mostremos que

$$I_1 = I_2. \tag{48.4}$$

Para todo elemento fijado G_k de la primera sucesión existe un número $k_0 = k_0(k)$ tal que

$$\bar{G}_k \subset D_{k_0}. \tag{48.5}$$

Efectivamente, si no se encontrara el número indicado k_0, para todo $m = 1, 2, \ldots$ natural existiría un punto $x^{(m)} \in \bar{G}_k \setminus D_m$. El conjunto abierto G_k, siendo medible según Jordan, es acotado, razón por la cual su clausura $\bar{G}_k$ representa un conjunto acotado cerrado, esto es, un compacto. Por ser acotado el conjunto, la sucesión $\{x^{(m)}\}$ es también acotada y, por ende, de acuerdo con el teorema de Bolzano—Weierstrass (véase el p. 18.1, teorema 2), se puede separar de ella una subsucesión convergente $\{x^{(m_\nu)}\}$. Si $x^{(0)} = \lim_{\nu \to +\infty} x^{(m_\nu)}$, entonces, debido a que el conjunto $\bar{G}_k$ es cerrado, tenemos $x^{(0)} \in \bar{G}_k$, y por eso $x_0 \in G$. Mas, en este caso, debido a la propiedad 2 de las sucesiones que agotan de modo monótono un conjunto (véase la definición 1), se encontrará un número m_0 tal que $D_{m_0} \ni x^{(0)}$. Dado que D_{m_0} es un conjunto abierto, será un entorno del punto $x^{(0)}$ y, por consiguiente, contiene casi todos los puntos de la sucesión $\{x^{(m_\nu)}\}$ que converge hacia $x^{(0)}$. Designemos mediante ν_0 un número cualquiera tal que sea $m_{\nu_0} \geqslant m_0$ y $x^{(m_{\nu_0})} \in D_{m_0}$. En virtud de la propiedad 1 de las sucesiones que agotan de modo monótono un conjunto, $x^{(m_{\nu_0})} \in D_{m_{\nu_0}}$, pero, como $x^{(m_{\nu_0})} \in \bar{G}_k$, esto contradice a la elección de la sucesión $\{x^{(m)}\}$.

De este modo queda demostrada la existencia del número k_0, mencionado anteriormente (véase (48.5)) (además, su existencia proviene también directamente del lema de Heine — Borel, véase el p. 18.3, puesto que el sistema $\{D_k\}$ forma un recubrimiento abierto del compacto $\bar{G}_k$).

Demos a conocer ahora que en virtud de la condición $f \geqslant 0$, de la inclusión (48.5) se desprende que $\int f dG_k \leqslant \int f dD_{k_0}$. Pero, evidentemente, $\int f dD_{k_0} \leqslant I_2$, por lo cual, para todo $k = 1, 2, \ldots$

$$\int f dG_k \leqslant I_2.$$

Pasando al límite en esta igualdad, para $k \to \infty$, obtendremos $I_1 \leqslant I_2$.

Del mismo modo se demuestra también la desigualdad $I_1 \geqslant I_2$. □

Ejemplo. Examinemos la integral $I = \iint\limits_{R^2} e^{-x^2-y^2} dxdy$. Pongamos $G_k = \{(x, y) : x^2 + y^2 < k^2\}$, $k = 1, 2, \ldots$. Ésta es una sucesión de conjuntos abiertos cuadrables (en el caso dado, simplemente una sucesión de círculos) que agota de módo monótono todo el plano R^2.

Sea $I_k = \iint\limits_{G_k} e^{-x^2-y^2} dxdy$. Pasemos a las coordenadas polares:

$$I_k = \iint\limits_{G_k} e^{-r^2} r dr d\varphi = \int\limits_0^{2\pi} d\varphi \int\limits_0^k -e^{-r^2} r dr = 2\pi \frac{e^{-r^2}}{2}\Bigg|_0^k = \pi(1 - e^{-k^2}).$$

De aquí, de acuerdo con la definición (48.1),

$$I = \lim_{k \to \infty} I_k = \pi. \tag{48.6}$$

La fórmula (48.6) permite hallar el valor de la integral

$$\int_{-\infty}^{+\infty} e^{-x^2} dx,$$

llamada *integral de Poisson* *) que se encuentra a menudo en diferentes aplicaciones. En efecto, designando con D_k el cuadrado $|x| \leqslant k$, $|y| \leqslant k$, $k = 1, 2, \ldots$, y aplicando a la integral, extendida a D_k, de la función $e^{-x^2-y^2}$ la fórmula de reducción de la integral múltiple a una reiterada (véase el p. 45.1), obtendremos

$$I = \iint_{R^2} e^{-x^2-y^2} dx dy = \lim_{k \to \infty} \iint_{D_k} e^{-x^2-y^2} dx dy =$$

$$= \lim_{k \to \infty} \int_{-k}^{k} dx \int_{-k}^{k} e^{-x^2-y^2} dx dy = \lim_{k \to \infty} \int_{-k}^{k} \left(\int_{-k}^{k} e^{-x^2} dx \right) e^{-y^2} dy = \cdot$$

$$= \lim_{k \to \infty} \left(\int_{-k}^{k} e^{-x^2} dx \right)^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right)^2.$$

Por esto de (48.6) se infiere inmediatamente que

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Teorema 2 (criterio de comparación). *Supongamos que en el conjunto abierto G se verifican las desigualdades $0 \leqslant f(x) \leqslant g(x)$, $x \in G$. En este caso la convergencia de la integral $\int f(x) dG$ se deduce de la convergencia de la integral $\int g(x) dG$, y la divergencia de la integral $\int g(x) dG$ se deduce de la divergencia de la integral $\int f(x) dG$.*

El teorema citado se demuestra por analogía con el teorema semejante para el caso unidimensional (véase el p. 33.3).

A título de ejemplos y patrones para la comparación con otras integrales consideraremos las siguientes

$$\int \cdots \int_{x_1^2 + \ldots + x_n^2 \leqslant 1} \frac{dx_1 \ldots dx_n}{(\sqrt{x_1^2 + \ldots + x_n^2})^{\alpha}}, \tag{48.7}$$

*) S. D. Poisson (1781—1840), matemático y físico francés.

$$\int\limits_{x_1^2+\ldots+x_n^2\leqslant 1}\cdots\int \frac{dx_1\ldots dx_n}{(\sqrt{x_1^2+\ldots+x_n^2})^{\alpha}}. \tag{48.8}$$

La primera integral se toma por el exterior de una bola unidad; la segunda, por el interior de la misma.

Con el fin de investigar estas integrales resulta cómodo introducir las *coordenadas esféricas* $\rho, \varphi_1, \ldots, \varphi_{n-1}$ en el espacio n-dimensional, sirviéndose de las fórmulas

$$\left.\begin{aligned}
x_1 &= \rho\cos\varphi_{n-1}\cos\varphi_{n-2}\cdots\cos\varphi_2\cos\varphi_1,\\
x_2 &= \rho\cos\varphi_{n-1}\cos\varphi_{n-2}\cdots\cos\varphi_2\operatorname{sen}\varphi_1,\\
x_3 &= \rho\cos\varphi_{n-1}\cos\varphi_{n-2}\cdots\cos\varphi_3\operatorname{sen}\varphi_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_i &= \rho\cos\varphi_{n-1}\cdots\cos\varphi_i\operatorname{sen}\varphi_{i-1},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_n &= \rho\operatorname{sen}\varphi_{n-1},
\end{aligned}\right\} \tag{48.9}$$

donde

$$0\leqslant\rho<+\infty,\quad 0\leqslant\varphi_1\leqslant 2\pi,\quad -\frac{\pi}{2}\leqslant\varphi_i\leqslant\frac{\pi}{2},\quad i=2,3,\ldots,n-1.$$

Con auyda de estas fórmulas a las coordenadas cartesianas $x_1, \ldots, x_n$ de un punto del espacio se las asignan las coordenadas esféricas $\rho, \varphi_1, \ldots, \varphi_{n-1}$, y viceversa. Con ello conviene tener en cuenta que, al igual que en el caso de las coordenadas polares en un plano, aquí no existe una correspondencia biunívoca completa entre los conjuntos de n números $(x_1, \ldots, x_n)$ y $(\rho, \varphi_1, \ldots, \varphi_{n-1})$.

Observemos que $\rho=\sqrt{x_1^2+\ldots+x_n^2}$.

Mediante cálculos elementales, pero un tanto engorrosos (que aquí no se realizan), podemos mostrar que el jacobiano de esta transformación tiene por expresión

$$\frac{\partial(x_1,x_2,\ldots,x_n)}{\partial(\rho,\varphi_1,\ldots,\varphi_{n-1})}=\rho^{n-1}\cos\varphi_2\cos^2\varphi_3\ldots\cos^{n-2}\varphi_{n-1}.$$

Pongamos, para abreviar,

$$\Phi(\varphi_2,\ldots,\varphi_{n-1})=\cos\varphi_2\cos^2\varphi_3\ldots\cos^{n-2}\varphi_{n-1}.$$

No es difícil convencerse de que $\Phi(\varphi_2, \ldots, \varphi_{n-1})\geqslant 0$ y que

$$c\overset{\text{def}}{=}\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\cdots\int_{-\pi/2}^{\pi/2}\Phi(\varphi_2\ldots\varphi_{n-1})d\varphi_1\ldots d\varphi_{n-1}>0.$$

Esto proviene inmediatamente de la propiedad 9 de las integrales múltiples citada en el p. 44.6.

Investiguemos ahora la convergencia de la integral (48.7). A título de sucesión de conjuntos medibles abiertos G_k, $k = 1, 2, \ldots$, que agota de modo monótono el exterior de la bola unidad Q, tomemos la sucesión de conjuntos

$$G_k = \{x = (\rho, \varphi_1, \ldots, \varphi_{n-1}) : 1 + \frac{1}{k} < \rho < k\}, \quad k = 1, 2, \ldots .$$

Pasemos a las coordenadas esféricas

$$\int\limits_{G_k} \ldots \int \frac{dx_1 \ldots dx_n}{(\sqrt{x_1^2 + \ldots + x_n^2})^\alpha} = \int\limits_{1+\frac{1}{k}}^{k} \int\limits_{0}^{2\pi} \int\limits_{-\pi/2}^{\pi/2} \ldots$$

$$\ldots \int\limits_{-\pi/2}^{\pi/2} \rho^{n-1-\alpha} \Phi(\varphi_2, \ldots, \varphi_{n-1}) d\rho d\varphi_1 \ldots d\varphi_{n-1} = c \int\limits_{1+\frac{1}{k}}^{k} \rho^{n-1-\alpha} d\rho .$$

De este modo, el problema de convergencia de la integral (48.7) se ha reducido a la convergencia de la integral $\int\limits_1^\infty \rho^{n-1-\alpha} d\rho$, la cual, como se sabe (véase el p. 33.3) converge cuando $n - 1 - \alpha < -1$, es decir, cuando $\alpha > n$, y diverge para $\alpha \leqslant n$. Así pues, queda demostrado el lema.

Lema 1. *La integral* (48.7) *converge, si* α *es superior a la dimensión del espacio y diverge en el caso contrario.*

Consideraremos ahora la integral (48.8). Al poner

$$G_k = \left\{x = (\rho, \varphi_1, \ldots, \varphi_{n-1}) : \frac{1}{k} < \rho < 1 - \frac{1}{k}\right\}, \quad k = 3, 4, \ldots,$$

obtendremos

$$\int\limits_{G_k} \ldots \int \frac{dx_1 \ldots dx_n}{(\sqrt{x_1^2 + \ldots + x_n^2})^\alpha} = \int\limits_{\frac{1}{k}}^{1-\frac{1}{k}} \int\limits_{0}^{2\pi} \int\limits_{\frac{\pi}{2}}^{\frac{\pi}{2}} \ldots$$

$$\ldots \int\limits_{-\pi/2}^{\pi/2} \rho^{n-1-\alpha} \Phi(\varphi_2, \ldots, \varphi_{n-1}) d\rho d\varphi_1 \ldots d\varphi_{n-1} = c \int\limits_{\frac{1}{k}}^{1-\frac{1}{k}} \rho^{n-1-\alpha} d\rho .$$

De este modo, el problema de convergencia de la integral (48.8) se ha reducido a la convergencia de la integral $\int\limits_0^1 \rho^{n-1-\alpha} d\rho$. Esta integral, como se sabe, converge, si $n - 1 - \alpha > -1$, es decir, si $\alpha < n$, y diverge en el caso contrario. El resultado obtenido enunciemos nuevamente en forma de un lema.

Lema 2. *La integral* (48.8) *converge, si* α *es inferior a la dimensión del espacio y diverge en el caso contrario.*

Al igual que en el caso unidimensional (véase el p. 33.3), sirviéndonos de las integrales (48.7) y (48.8), podemos enunciar los criterios de convergencia de las integrales múltiples impropias, sin embargo aquí no nos detendremos en esto detalladamente.

48.3. INTEGRALES IMPROPIAS DE LAS FUNCIONES QUE CAMBIAN DE SIGNO

Definición 3. *La integral impropia* $\int f dG$ *se llama absolutamente convergente, si converge la integral* $\int |f| dG$.

Para estudiar la convergencia absoluta de una integral de la función $f(x)$, nos serán útiles las funciones

$$f_+(x) = \begin{cases} f(x), & \text{si } f(x) \geqslant 0, \\ 0, & \text{si } f(x) < 0 \end{cases} \qquad f_-(x) = \begin{cases} -f(x), & \text{si } f(x) \leqslant 0, \\ 0, & \text{si } f(x) > 0 \end{cases}$$

Es fácil ver que

$$f_+ = \frac{|f| + f}{2}, \quad f_- = \frac{|f| - f}{2}, \tag{48.10}$$

$$0 \leqslant f_+(x) \leqslant |f(x)|, \quad 0 \leqslant f_-(x) \leqslant |f(x)|, \tag{48.11}$$

$$f(x) = f_+(x) - f_-(x), \quad |f(x)| = f_+(x) + f_-(x). \tag{48.12}$$

De las fórmulas (48.10) se deduce que si la función f es integrable según Riemann en cierta región medible según Jorgan, las funciones f_+ y f_- serán también integrables según Riemann en dicha región; de la primera fórmula (48.12) proviene la afirmación recíproca. Por ello, de (48.10) — (48.12) se infiere que la *integral* $\int f dG$ *converge absolutamente cuando, y sólo cuando, convergen las integrales*

$$\int f_+ dG \ \text{ y } \ \int f_- dG.$$

Al igual qué en el caso de integrales impropias de la función de una sola variable, de la convergencia absoluta de una integral múltiple proviene la convergencia de la misma (siempre que, por supuesto, se consideran sólo unas funciones que son integrables en cada conjunto medible abierto, contenido junto con su clausura en un conjunto abierto por el cual se realiza la integración). Esto proviene en seguida de las fórmulas (48.11), la primera fórmula (48.12) y del teorema 2 del presente párrafo (véase el p. 48.2). No obstante, para las integrales múltiples propias es válido también el teorema recíproco.

Teorema 3. *Si una integral múltiple* $\int f dG$ $(n \geqslant 2)$ *converge, converge absolutamente.*

Este teorema, algo sorprendente a primera vista, está relacionado con la diferencia en la definición de las integrales impropias de la función de una variable y de n variables ($n > 1$), citadas al principio de este párrafo *).

*) Hemos de notar, sin embargo, que en el caso n-dimensional también podríamos obtener la misma relación entre la convergencia y la convergencia absoluta de una integral que existe en el caso unidimensional, siempre que se introduzca del modo correspondiente la defi-

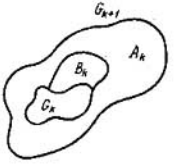


Fig. 211

DEMOSTRACIÓN DEL TEOREMA. Supongamos que la integral $\int f dG$ diverge absolutamente, es decir, para cierta sucesión (y, por lo tanto, para cualquier sucesión véase el teorema 1 en el p. 48.2) de conjuntos abiertos medibles según Jordan G_k, $k = 1, 2, \ldots$, que agota de modo monótono el conjunto abierto G, se tiene $\lim\limits_{k\to\infty} \int |f| dG_k = +\infty$. Sin restringir la generalidad de razonamientos (pasando, si es necesario, a una subsucesión) podemos suponer que

$$\int |f| dG_{k+1} > 3 \int |f| dG_k + 2k, \quad k = 1, 2, \ldots . \tag{48.13}$$

Sea $A_k = G_{k+1} \setminus \bar{G}_k$; en este caso A_k es un conjunto medible abierto y, como $\bar{G}_k \subset G_{k+1}$, entonces (fig. 211) $G_{k+1} = A_k \cup \bar{G}_k$, y, por consiguiente,

$$\int |f| dG_{k+1} = \int |f| dA_k + \int |f| dG_k.$$

De aquí, en vista de la desigualdad (48.13), $\int |f| dA_k > 2 \int |f| dG_k + 2k$. Haciendo uso de la segunda fórmula (48.12), obtendremos

$$\int f_+ dA_k + \int f_- dA_k > 2 \int |f| dG_k + 2k.$$

Sea, para concretar, $\int f_+ dA_k \geqslant \int f_- dA_k$; en este caso

$$2 \int f_+ dA_k \geqslant \int f_+ dA_k + \int f_- dA_k > 2 \int |f| dG_k + 2k,$$

y, por lo tanto,

$$\int f_+ dA_k > \int |f| dG_k + k. \tag{48.14}$$

Nuestro objetivo es obtener una desigualdad del tipo semejante no para la función f_+, sino para la función f. Con este fin parece que podríamos omitir simplemente todos los puntos, donde la función f_+ se anula; entonces en el conjunto que

---

nición de la integral n-múltiple impropia. Por ejemlo, en el caso de las integrales extendidas a todo el espacio, resulta suficiente para ello que en la definición de la integral a título de elementos de la sucesión, que agota de modo monótono un conjunto, se tomen sólo las bolas n-dimensionales con centro en el origen de coordenadas. Por otra parte, aplicando a una integral unidimensional la definición de integral impropia citada en el p. 48.1 y entendiendo la integral unidimensional de Riemann en el sentido del §. 44, el teorema 3 y su demostración tienen lugar también para $n = 1$.

queda tendríamos $f = f_+$. No obstante, el conjunto obtenido puede, en el caso general, resultar no medible, razón por la cual recurriremos al rodeo.

De la desigualdad (48.14) proviene que realizada cualquier partición suficientemente menuda $\tau = \{X_i\}_{i=1}^{i=i_0}$ del conjunto A_k (véase el p. 44.3), para cualquier suma integral de Riemann tenemos

$$\sum_{i=1}^{i_0} f_+(\xi_i)\mu X_i > \int |f| dG_k + k, \quad \xi_i \in X_i, \quad i = 1, 2, \ldots, i_0.$$

Elijamos la partición mencionada τ del conjunto medible abierto A_k de una manera tal que todos los elementos X_i de dicha partición sean también conjuntos abiertos medibles según Jordan, lo que es siempre posible (véase el p. 44.4). Designemos con X_i^* aquellos conjuntos $X_i \in \tau$, para los cuales $f_+(\xi) > 0$ en todos los puntos $\xi \in \overline{X}_i$; en este caso, al elegir $\xi_i \in \overline{X}_i \neq \overline{X}_i^*$ de tal modo que sea $f(\xi_i) = 0$, obtendremos

$$\sum_i{}' f_+(\xi_i)\mu X_i^* > \int |f| dG_k + k, \tag{48.15}$$

donde (como también en lo sucesivo) la "virgulilla" de la suma significa que la sumación se extiende sólo a aquellos índices i, para los cuales $X_i = X_i^*$. Pongamos $B_k = \bigcup' X_i^*$ (véase fig. 211). Es obvio que B_k es un conjunto abierto medible dispuesto dentro del conjunto A_k, y $\tau^* = \{X_i^*\}$ constituye su partición. En la clausura de este conjunto $f_+ > 0$, y, por consiguiente, $f_+ = f$. De la desigualdad (48.15) se deduce que para la suma inferior de Darboux s_{τ^*} de la función f en el conjunto B_k se verifica la desigualdad $s_{\tau^*} \geqslant \int f dG_k + k$. De aquí se desprende, evidentemente, que

$$\int f dB_k \geqslant \int |f| dG_k + k. \tag{48.16}$$

Observemos que $f \geqslant -|f|$, y, por ende,

$$\int f dG_k \geqslant -\int |f| dG_k. \tag{48.17}$$

Al sumar las desigualdades (48.16) y (48.17), obtendremos:

$$\int f dB_k + \int f dG_k \geqslant k. \tag{48.18}$$

Sea $D_k = B_k \cup G_k$, $k = 1, 2, \ldots$. Es evidente que D_k es un conjunto abierto medible y

$$G_k \subset D_k \subset G_{k+1}, \quad k = 1, 2, \ldots. \tag{48.19}$$

En vista de que los conjuntos B_k y G_k no se intersecan (puesto que no se intersecan los conjuntos A_k y G_k), de (48.18) tenemos

$$\int f dD_k \geqslant k,$$

de donde

$$\lim_{k \to \infty} \int f dD_k = +\infty. \tag{48.20}$$

De la inclusión (48.19) se infiere que los conjuntos D_k, $k = 1, 2, \ldots$, forman una sucesión de conjuntos abiertos medibles, que agota de modo monótono el conjunto abierto G, pues tal era la sucesión dada G_k, $k = 1, 2, \ldots$, por lo cual la igualdad (48.20) significa que la integral $\int f dG$ diverge. □

Así pues, para las integrales múltiples, la convergencia de la integral impropia $\int f dG$ es equivalente a la convergencia absoluta de ésta.

Ejercicio 2. Al sustituir en la definición de la integral múltiple impropia los conjuntos abiertos en todos los puntos por unas regiones (en particular, considerando solamente las sucesiones, compuestas sólo de las regiones medibles, que agotan de modo monótono la región dada) muéstrese que aún para tal definición "más estrecha" de la integral múltiple impropia queda en vigor el teorema 3.

§ 49. ALGUNAS APLICACIONES GEOMÉTRICAS Y FÍSICAS DE LAS INTEGRALES MÚLTIPLES

49.1. CÁLCULO DE ÁREAS Y DE VOLÚMENES

Sea X un conjunto medible en R^n. Como se sabe (véase el p. 44.6):

$$\mu X = \int dX. \tag{49.1}$$

De este modo, con ayuda de la integral n-múltiple podemos calcular la medida de los conjuntos medibles en un espacio n-dimensional (un área en el espacio bidimensional, un volumen, en el espacio tridimensional). Si la integral n-múltiple (49.1) puede reducirse a la reiterada (véase § 45), el cálculo de la medida del conjunto medible X de un espacio n-dimensional se reducirá al cálculo de una integral $(n-1)$-múltiple.

Sea, por ejemplo, D un conjunto abierto medible en el espacio $(n-1)$-dimensional $R^{n-1}_{x_1}, \ldots, x_{n-1}$, siendo $x_n = f(x_1, \ldots, x_{n-1})$ una función no negativa definida y continua en la clausura $\bar{D}$ del conjunto D, y

$$G = \{x = (x_1, \ldots, x_n) : (x_1, \ldots, x_{n-1}) \in D, \quad 0 < x_n < f(x_1, \ldots, x_{n-1})\}$$

(de este modo, G representa un análogo n-dimensional del trapecio plano curvilíneo examinado en el p. 32.1). En este caso

$$\mu G = \int dG = \int dD \int_0^{f(x_1, \ldots, x_{n-1})} dx_n = \int f(x_1, \ldots, x_{n-1}) dD,$$

es decir,

$$\mu G = \int_D \cdots \int f(x_1, \ldots, x_{n-1}) dx_1 \ldots dx_{n-1}.$$

La medida de los conjuntos abiertos del espacio R^n, $n \geqslant 2$, arbitrarios (no forzosamente medibles según Jordan) y, en particular, no acotados, si ella se entiende en el sentido de la definición del p. 31.1 y 31.2, es decir, como una medida inferior de

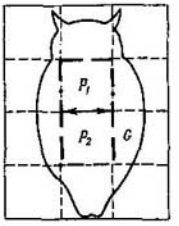


Fig. 212

Jordan, puede ser calculada mediante las integrales impropias. En efecto, sea G un conjunto abierto arbitrario en R^n y sea G_k, $k = 1, 2, \ldots$, una sucesión de conjuntos abiertos medibles que agota de modo monótono el conjunto G (véase el p. 48.1). Entonces, como se sabe, (véase el p. 31.2), $\lim\limits_{k\to\infty} \mu G_k = \mu G$. Mas, en virtud de (49.1), $\mu G_k = \int dG_k$, por lo cual $\mu G = \lim\limits_{k\to\infty} \int dG_k$.

Según la definición de la integral múltiple impropia, $\lim\limits_{k\to\infty} \int dG_k = \int dG$. De este modo, $\mu G = \int fdG$, donde la integral en el segundo miembro se entiende, en el caso general (a saber, si a ciencia cierta G no es una región medible) como impropia.

Queda sólo mostrar que para cualquier conjunto abierto G existe siempre una sucesión de conjuntos medibles G_k, $k = 1, 2, \ldots$, que agotan de modo monótono el conjunto dado G. Demostrémoslo.

Consideraremos la sucesión T_k, $k = 1, 2, \ldots$, de particiones del espacio R^n en cubos (véase el p. 44.1) y designemos mediante Q_k un cubo abierto n-dimensional, determinado de la manera siguiente.

$$Q_k = \{(x_1, \ldots, x_n) : |x_i| < k,\ i = 1, 2, \ldots, n\}.$$

El número de los cubos de rango dado k (véase el p. 44.1), contenidos en Q_k, y, por consiguiente, con la mayor razón, en la intersección $G \cap Q_k$, es finito. Designemos estos cubos cerrados $P_1, \ldots, P_{j_k}$:

$$P_j \in T_k; \quad P_j \subset G \cap Q_k, \quad j = 1, 2, \ldots, j_k.$$

Con G_k se designará el conjunto de puntos interiores del conjunto $\bigcup\limits_{j=1}^{j_k} P_j$. Por ejemplo, en el caso que se expone en la fig. 212, el conjunto G_k se compone de los puntos interiores de dos cuadrados, P_1 y P_2, y un intervalo que se obtiene extrayendo los vértices de estos cuadrados de su arista común.

Precisamente los conjuntos G_k, $k = 1, 2, \ldots$, son conjuntos abiertos medibles que forman la sucesión que agota de modo monótono el conjunto abierto dado G.

Recordemos que para calcular los volúmenes de los cuerpos resulta, a menudo, cómodo el método de secciones: véase la fórmula (45.23).

Ejercicio 1. Demuéstrese que la sucesión construida de conjuntos G_k, $k = 1, 2, \ldots$, forma realmente una sucesión de conjuntos medibles que agota de modo monótono el conjunto dado G.

49.2. APLICACIONES FÍSICAS DE LAS INTEGRALES MÚLTIPLES

Con ayuda de las integrales múltiples se pueden calcular diferentes magnitudes físicas: masa y carga de un cuerpo, centro de gravedad, momento de inercia, flujo de líquido, potencial de un cuerpo, etc.

Hallemos, como un ejemplo, el centro de gravedad de una figura plana. Supongamos que en una cierta región cuadrable G está distribuida cierta masa cuya densidad superficial $\rho(x, y)$ es, en el caso general, variable, es decir, en la clausura $\bar{G}$ de la región G está definida una función continua y no negativa $\rho(x, y)$. La región G con la masa distribuida en ella se llamará *figura* S, y la magnitud

$$M = \iint_G \rho(x, y)dxdy \tag{49.2}$$

la *masa* de dicha figura. Si $\rho(x, y)$ no es cero idéntico, $M > 0$.

Determinemos y hallemos el centro de gravedad de la figura S. Sea $\tau = \{G_i\}$, $i = 1, 2, \ldots k$, una partición de la región G (véase el p. 44.3). Se denominará figura S_i un conjunto G_i con la masa distribuida en él de densidad $\rho(x, y)$, $(x, y) \in G_i$. Elijamos para cada i el punto correspondiente $(\xi_i, \eta_i) \in \bar{G}_i$. La magnitud $m_i = \rho(\xi_i, \eta_i)\mu G_i$ se llamará valor aproximado de la masa de la figura S_i (tal denominación es natural según (49.2)). Entre tanto, las magnitudes $m_i\xi_i$ y $m_i\eta_i$ se denominarán valores aproximados de los momentos estáticos de la figura S_i, $i = 1, 2, \ldots, k$, respecto de los ejes coordenados Oy y Ox, respectivamente (es natural llamarlas así, porque las magnitudes my y mx (véase el p. 32.6) llevan el nombre de momentos estáticos de un punto material de masa m con coordenadas (x, y) respecto de los ejes Ox y Oy). Por fin, las magnitudes

$$\begin{aligned} S_x(\tau) &= \sum_{i=1}^{k} \eta_i m_i = \sum_{i=1}^{k} \eta_i \rho(\xi_i, \eta_i)\mu G_i, \\ S_y(\tau) &= \sum_{i=1}^{k} \xi_i m_i = \sum_{i=1}^{k} \xi_i \rho(\xi_i, \eta_i)\mu G_i, \end{aligned} \tag{49.3}$$

se llamarán τ-momentos aproximados de la figura S con respecto a los ejes Ox y Oy, mientras que sus límites, para $\delta_\tau \to 0$,

$$\lim_{\delta_\tau \to 0} S_x(\tau) = S_x, \qquad \lim_{\delta_\tau \to 0} S_y(\tau) = S_y,$$

momentos estáticos de la figura S respecto de los ejes Ox y Oy. Bajo las suposiciones asumidas, estos límites existen. Efectivamente, de las fórmulas (49.3) se ve que $S_x(\tau)$ y $S_y(\tau)$ son las sumas integrales de Riemann para las funciones $y\rho(x, y)$ y

$x\rho(x, y)$ y por esta razón

$$S_x = \iint\limits_G y\rho(x, y)dxdy, \quad S_y = \iint\limits_G x\rho(x, y)dxdy. \tag{49.4}$$

Definición 1. *El punto (x_0, y_0) se denomina centro de gravedad (centro de masa, centro de inercia) de la figura S, si los momentos estáticos respecto de los ejes de coordenadas del punto material de masa M, que equivale a la masa de toda la figura S y que se encuentra en el punto (x_0, y_0), son iguales a los momentos estáticos correspondientes de la figura S, es decir, si*

$$Mx_0 = S_y, \quad My_0 = S_x.$$

De las fórmulas (49.2) y (49.3) obtenemos

$$x_0 = \frac{\iint\limits_G x\rho(x, y)dxdy}{\iint\limits_G \rho(x, y)dxdy}, \quad y_0 = \frac{\iint\limits_G y\rho(x, y)dxdy}{\iint\limits_G \rho(x, y)dxdy}.$$

Ejercicio 2. Demuéstrese que el centro de gravedad de una figura no depende de cómo se escoge el sistema de coordenadas.

A título de ejemplo consideraremos un "trapecio curvilíneo" G engendrado por las gráficas de las funciones continuas no negativas $f(x)$ y $g(x)$, $0 \leqslant g(x) \leqslant f(x)$, $a \leqslant x \leqslant b$:

$$G = \{(x, y) : a < x < b, \quad g(x) < y < f(x)\}.$$

Sea $\rho(x, y) = 1$. Dado que $\iint\limits_G dxdy = \mu G$, se tiene

$$x_0 = \frac{1}{\mu G} \iint\limits_G xdxdy = \frac{1}{\mu G}\int\limits_a^b xdx \int\limits_{g(x)}^{f(x)} dy = \frac{1}{\mu G}\int\limits_a^b [f(x) - g(x)]xdx,$$

$$y_0 = \frac{1}{\mu G} \iint\limits_G ydxdy = \frac{1}{\mu G}\int\limits_a^b dx \int\limits_{g(x)}^{f(x)} ydy = \frac{1}{2\mu G}\int\limits_a^b [f^2(x) - g^2(x)]dx;$$

de aquí

$$2\pi y_0 \mu G = \pi \int\limits_a^b f^2(x)dx - \pi \int\limits_a^b g^2(x)dx.$$

Aquí en el segundo miembro interviene el volumen de un cuerpo obtenido por revolución del trapecio curvilíneo G en torno al eje de x, es decir, hemos llegado al *segundo teorema de Guldin.*

Teorema (de Guldin). *El volumen de un cuerpo de revolución de una figura plana alrededor de un eje que no la corta es igual al producto del área de esta figura por la longitud de la circunferencia descrita por el centro de gravedad de la misma.*

Ejemplo. Sirviéndose del segundo teorema de Guldin, calcúlese el volumen μQ del toro Q, obtenido por revolución del círculo $(x-a)^2+y^2 \leqslant r^2$, $0<r \leqslant a$, alrededor del eje Oy:

$$\mu Q = 2\pi a \cdot \pi r^2 = 2\pi^2 a r^2.$$

Ejercicios. Hállese la masa de una figura plana limitada por las líneas:

3. $y^2 = 2x$, $x+y=4$, $y \geqslant 1$; $\rho(x, y) = x+y$.
4. $y = 2x$, $y = -2$, $y = 4x-2$; $\rho(x, y) = 2|x| + |y|$.

Hállense los momentos estáticos respecto de los ejes de coordenadas de la figura plana homogénea ($\rho = \rho_0 = \text{const}$), limitada por las líneas dadas:

5. $y^2 = 4x$, $x+y=3$, $x=0$.
6. $y = x^3$, $x+y=2$, $x=0$.

Hállense las coordenadas del centro de masas de una figura plana limitada por las líneas indicadas:

7. $x^2+y^2=4$, $x^2+y^2=1$, $y \geqslant 0$; $\rho = \rho_0 = \text{const}$.
8. $y^2 = 4x$, $y=2$, $x=0$; $\rho(x, y) = x$.
9. $r=\sqrt{2}$, $r = 2\operatorname{sen}\varphi$ $(0 \leqslant \varphi \leqslant \pi/2, r \geqslant \sqrt{2})$, $\rho = \rho_0 = \text{const}$ (r, φ son las coordenadas polares).
10. $x = a\cos^3 t$, $y = a\operatorname{sen}^3 t$, $y=0$ $(0 \leqslant t \leqslant \pi)$, $\rho = \rho_0 = \text{const}$.

§ 50. ELEMENTOS DE LA TEORÍA DE SUPERFICIES

50.1. CONCEPTO DE SUPERFICIE

Sea dado en el espacio R^3 un sistema fijo de coordenadas rectangulares cartesianas x, y, z. Las coordenadas cartesianas en los planos, donde se disponen las regiones que se aplican, se designarán con las letras u, v; las propias regiones, con la letra D y las aplicaciones de éstas que han de ser examinadas, con letras f, r, ρ (quizás, acompañadas de tales o cuales índices). Como hasta ahora, con $\bar{D}$ se denotará la clausura de la región D (recordemos que $\bar{D}$ se denomina región cerrada) y mediante ∂D, la frontera de D (véase el p. 18.2). Para las imágenes de los puntos $M=(u, v) \in \bar{D}$ se emplearán, en las aplicaciones mencionadas, tanto la inscripción del tipo $f(M)$, como la del tipo $f(u, v)$.

Se denomina *superficie continua* S todo conjunto de puntos del espacio tridimensional R^3 dado como una imagen continua de cierta región plana cerrada $\bar{D}$. La propia aplicación continua $r(u, v)$ de la región cerrada $\bar{D}$ sobre el conjunto S se llama *representación de la superficie* (o, más detalladamente, *representación paramétrica*) y se escribe

$$S = \{r(u, v) : (u, v) \in \bar{D}\}.$$

Las variables u, v llevan el nombre de *coordenadas o parámetros* de la superficie continua S.

Para la superficie continua $S = \{r = r(u, v) : (u, v) \in \bar{D}\}$ el conjunto de puntos del espacio R^3, dado como una imagen de la frontera ∂D de la región D en la aplicación $r(u, v)$, se denomina *borde de la superficie* S y se designa por ∂S:

$$\partial S = \{r(u, v) : (u, v) \in \partial D\}.$$

Por analogía con la definición de una curva podemos introducir el concepto de aplicaciones equivalentes, pero esta vez no de los segmentos, sino de las regiones planas cerradas en el espacio tridimensional R^3 y considerar, según la definición, que dos aplicaciones continuas equivalentes prefijan una misma superficie continua (véase el p. 50.2*). Las aplicaciones que realizan la equivalencia de dos representaciones de una misma superficie se denominan *transformaciones admisibles de los parámetros*.

Dada la representación $r(u, v)$, $(u, v) \in D$, de cierta superficie continua y fijados los valores de los parámetros u, v, resulta natural que mediante $r(u, v)$ se designe un punto de dicha superficie en el que se aplica, para la representación en consideración, el punto $(u, v) \in \bar{D}$.

Subrayemos que la representación de una superficie continua no es forzosamente la aplicación biunívoca. Un punto de la superficie continua $S = \{r(u, v); (u, v) \in \bar{D}\}$, en el que, realizándose la aplicación dada $r(u, v)$, se aplican por lo menos dos puntos diferentes de la región cerrada $\bar{D}$, se llama *punto múltiple* de dicha superficie.

De este modo, si el punto M de una superficie continua es el punto múltiple de la misma, entonces, dada la representación $r(u, v)$, $(u, v) \in \bar{D}$, de esta superficie, existen por lo menos dos puntos $(u_1, v_1) \in \bar{D}$ y $(u_2, v_2) \in \bar{D}$ tales que $r(u_1, v_1) = r(u_2, v_2) = M$.

La aplicación $r(u, v)$ puede ser definida en la forma coordenada:

$$r(u, v) = (x(u, v), y(u, v), z(u, v))$$

y en la vectorial:

$$r = r(u, v),$$

donde $r(u, v)$ es el radio vector con extremo en el punto $r(u, v) \in R^3$.

En lo que sigue se estudiarán, ante todo, las propiedadas diferenciales de las superficies de ciertas clases que se componen de superficies "suficientemente suaves", es decir, de aquellas superficies que son derivables (continuamente) un número suficiente de veces. Demos a conocer, por ejemplo, la noción de superficie continuamente derivable.

Se denomina *superficie continuamente derivable* un conjunto S del espacio R^3 dado como una imagen continuamente derivable de cierta región plana cerrada.

La propia aplicación continuamente derivable de la región cerrada $\bar{D}$ sobre el conjunto S se llama, como hasta ahora, representación de esta superficie, con la particularidad de que se considera, según la definición, que dos aplicaciones continuamente derivables de las regiones planas cerradas determinan una misma superficie continuamente derivable, siempre que son equivalentes respecto de las aplicaciones continuamente derivables (véase el p. 50.2*).

Del modo análogo se determinan también otras clases especiales de las superficies continuas: las superficies dos veces continuamente derivables y, en general, las superficies n veces continuamente derivables.

Si en una de las representaciones de una superficie continua podemos escoger como parámetros dos coordenadas cualesquiera del espacio R^3 (por ejemplo, si existen una región cerrada D en el plano xy y una función $z = f(x, y)$, $(x, y) \in \bar{D}$,

que es la representación de la superficie continua en consideración), entonces tal representación se llama explícita.

Evidentemente, si una superficie continua admite la representación explícita, está privada de puntos múltiples.

En adelante una superficie continua se llamará simplemente superficie, siempre que esto no nos lleve a una equivocación.

Ejemplo. La superficie determinada por la representación

$$x = r\cos\psi\cos\varphi, \quad y = r\cos\psi\,\operatorname{sen}\varphi, \quad z = r\,\operatorname{sen}\psi,$$
$$0 \leqslant \varphi \leqslant 2\pi, \quad -\frac{\pi}{2} \leqslant \psi \leqslant \frac{\pi}{2},$$

es una esfera con centro en el origen de coordenadas y radio r; todo el meridiano de dicha esfera se compone de puntos múltiples.

En el punto que sigue daremos otra definición de la superficie que será, en cierto sentido, más detallada. Es conveniente, obviamente, omitir dicho punto en el transcurso de la primera lectura y volver a leerlo cuando se sienta la necesidad intrínseca.

50.2*. APLICACIONES EQUIVALENTES. SUPERFICIES DADAS EN FORMA PARAMÉTRICA

Para definir estrictamente una superficie es preciso, ante todo, introducir la noción de aplicaciones equivalentes de las regiones planas cerradas.

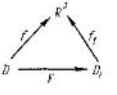
Definición 1. *Una aplicación continua f de la clausura $\overline{D}$ de cierta región plana D en el espacio tridimensional R^3 se denomina equivalente a la aplicación continua f_1 de la clausura $\overline{D}_1$ de la región plana D_1 en el mismo espacio R^3, si existe tal aplicación homeomorfa (véase la definición 5 en el p. 41.4) F de la región cerrada $\overline{D}$ sobre otra región cerrada $\overline{D}_1$ que, cuando se realiza dicha aplicación, los puntos interiores pasan a puntos interiores y los límites, también a puntos límites (es decir, D se aplica sobre D_1 y ∂D, sobre ∂D_1) y para todo punto $M \in D$ se verifica la igualdad*

$$f(M) = f_1[F(M)], \tag{50.1}$$

es decir, $f = f_1 \circ F$.

En este caso F recibe el nombre de *aplicación que realiza la equivalencia de las aplicaciones f y f_1*. Si f es equivalente a la aplicación f_1, se escribe $f \sim f_1$.

La definición de las aplicaciones equivalentes puede ser expresada en forma esquemática mediante un diagrama, donde las flechas significan las aplicaciones en consideración y el resultado de las aplicaciones no depende de la elección del trayecto en el diagrama:



Es obvio que:

1) toda aplicación es equivalente a sí misma: $f \sim f$ (aquí, la aplicación que realiza la equivalencia es aplicación idéntica).

Es fácil convencerse de que

2) si $f \sim f_1$, entonces $f_1 \sim f$,

3) si $f \sim f_1$ y $f_1 \sim f_2$, entonces $f \sim f_2$.

Si f y f_1 son las aplicaciones continuas equivalentes de las regiones cerradas respectivas $\overline{D}$ y $\overline{D}_1$, entonces de (50.1) proviene que las imágenes de los conjuntos $\overline{D}$ y $\overline{D}_1$ en las aplicaciones f y f_1 coinciden:

$$f(\overline{D}) = f_1(\overline{D}_1). \tag{50.2}$$

Diremos, además, que las condiciones impuestas sobre las aplicaciones equivalentes en la definición 1 son independientes. A saber, de lo que F es una aplicación homeomorfa de la región cerrada D sobre otra región cerrada $\overline{D}_1$ no se infiere que transforma los puntos interiores en interiores. Por ejemplo, si $D = \{(u, v) : u^2 + v^2 < 1\}$ es un círculo y $D_1 = \{(u, v) : 0 < u^2 + v^2 < 1\}$ es un círculo con centro "punzado", entonces la aplicación idéntica (que es, evidentemente, homeomorfa) de $\overline{D}$ sobre $\overline{D}_1$ hace pasar el punto interior (0, 0) de la región D al punto límite (0, 0) de la región D_1.

Demos ahora la definición de superficie.

Definición 2. *Cualquier conjunto de toda clase de aplicaciones continuas $r(u, v)$, equivalentes entre sí (véase la definición 1), de las regiones planas cerradas $\overline{D}$ en el espacio tridimensional R^3 se denomina superficie S dada en forma paramétrica y se designa*

$$S = \{r(u, v) : (u, v) \in \overline{D}\}, \tag{50.3}$$

y cada una de las aplicaciones continuas equivalentes citadas $r(u, v)$ lleva el nombre de representación de la superficie S dada en forma paramétrica.

Si $r(u, v)$, $(u, v) \in \overline{D}$, es una representación de la superficie S dada en forma paramétrica y si $r(u, v)$ es el radio vector con extremo en el punto $r(u, v)$, entonces $r(u, v)$, $(u, v) \in \overline{D}$, se llama *representación vectorial* de la superficie citada S y se escribe

$$S = \{r(u, v), (u, v) \in \overline{D}\}. \tag{50.4}$$

Si $r(u, v) = (x(u, v), y(u, v), z(u, v))$, las funciones

$$x = x(u, v), \quad y = y(u, v), \quad z = (u, v)$$

llevan el nombre de *representación coordenada* de la superficie S dada en forma paramétrica y se escribe en este caso

$$S = \{x(u, v), y(u, v), z(u, v); (u, v) \in \overline{D}\}. \tag{50.5}$$

Es evidente que una superficie dada en forma paramétrica se determina unívocamente por cada una de sus representaciones. Esto permite concebir (lo que a menudo resulta más conveniente) el segundo miembro de cada una de las igualdades (50.3), (50.4) y (50.5) no como una totalidad de todas las representaciones de tipo determinado de la superficie en consideración S, sino como su representación correspondiente bien determinada.

Definición 3. *Sean $r(M)$, $M \in \overline{D}$, y $\rho(M_1)$, $M_1 \in \overline{D}_1$, dos representaciones de cierta superficie S dada en forma paramétrica y sea F la aplicación de $\overline{D}$ sobre $\overline{D}_1$ que realiza su equivalencia (véase la definición 1).*

Si $M_1 = F(M)$, $M \in \bar{D}$, $M_1 \in \bar{D}_1$ (el punto M, y, por ende, el punto M_1 también están fijos) y, por consiguiente, $r(M) = \rho(M_1) = P \in R^3$, entonces los pares (P, M) y (P, M_1) se llaman equivalentes y se escribe en este caso

$$(P, M) \sim (P, M_1).$$

Es fácil comprobar que

1) $(P, M) \sim (P, M)$;

2) si $(P, M) \sim (P, M_1)$, entonces $(P, M_1) \sim (P, M)$;

3) si $(P, M) \sim (P, M_1)$, y $(P, M_1) \sim (P, M_2)$, entonces $(P, M) \sim (P, M_2)$.

Si $(P, M) \sim (P, M_1)$ y M es un punto interior (límite) de la región cerrada $\bar{D}$, entonces, de acuerdo con la definición 1, M_1 será también un punto interior (límite) de la región cerrada $\bar{D}_1$.

Definición 4. *Sea S una superficie dada en forma paramétrica. Cualquier totalidad $\{(P, M)\}$, $M \in \bar{D}$, de todos los pares equivalentes entre sí (P, M) (el punto $P \in R^3$ está fijo) se denomina punto de la superficie dada S y el punto P, su portador.*

Un punto $\{(P, M)\}$, $M \in D$, de la superficie S se llama interior (de contorno), si cada punto M es interior (de frontera) para la región cerrada correspondiente $\bar{D}$.

Todo punto $\{(P, M)\}$, $M \in \bar{D}$, de una superficie $S = \{r(M), M \in \bar{D}\}$, dada paramétricamente se determina unívocamente por cada par $(P, M) \in \{(P, M)\}$, y, como en este par $P = r(M)$, cualquier punto de la superficie S, definida en forma paramétrica, para cierta representación dada $r(M)$, $M \in \bar{D}$, de esta superficie, se determina unívocamente por el punto M, con la particularidad de que el punto $P = r(M)$ es el portador del punto considerado de la superficie. Por esta razón, para abreviar, los puntos de una superficie dada en forma paramétrica se designarán, por regla general, no mediante el símbolo $\{(P, M)\}$, sino simplemente $r(M)$, o bien, lo que es igual, $r(u, v)$, donde $M = (u, v)$. En virtud de lo expuesto, esta designación posee el significado unívoco.

Definición 5. *Una totalidad de todos los portadores de todos los puntos de la superficie S dada en forma paramétrica se denomina portador de dicha superficie.*

En vista de la condición (50.2), el portador de una superficie dada en forma paramétrica (el que representa, evidentemente, cierto conjunto de puntos en el espacio R^3) se determina unívocamente por cualquiera de sus representaciones.

Definición 6. *El punto $P \in R^3$ que sirve de portador de dos puntos distintos de la superficie S dada paramétricamente se llama punto múltiple del portador de la superficie dada en forma paramétrica.*

Volviendo a la definición de superficie, proporcionada en el p. 50.1, vemos que lo que se ha denominado en aquella circunstancia "*superficie continua*" se llama en términos nuevos "*portador de una superficie dada en forma paramétrica*". La tentativa de introducir la noción de "punto de la superficie" para las superficies provistas de puntos múltiples conduce en tal o cual forma a las definiciones 4 y 6. Recalquemos que la noción de superficie dada en forma paramétrica con puntos múltiples resulta muy cómoda para considerar toda una serie de problemas que se analizan en los párrafos que siguen.

En adelante, siempre que no surjan equivocaciones, la "superficie continua" (véase el p. 50.1) o bien, que es lo mismo, el "portador de una superficie dada para-

métricamente" (véase la definición 5), como también la "superficie dada paramétricamente" (véase la defición 2) se llamarán simplemente superficie.

Definamos ahora el concepto de una parte de la superficie.

Definición 7. *Supongamos que S representa una superficie dada paramétricamente; $r(u, v)$, $(u, v) \in \bar{D}$, es cierta representación de dicha superficie; D', una región contenida en D: $D' \subset D$. La superficie S' dada paramétricamente que se determina por la representación $r(u, v)$, $(u, v) \in \bar{D}$, se denomina parte de la superficie S.*

Como ya se ha indicado (véase el p. 50.1), el concepto de aplicaciones equivalentes de las regiones planas cerradas puede introducirse no sólo para las aplicaciones continuas, sino también para otras clases de aplicaciones, por ejemplo, para aquellas que son continuamente derivables. En el caso de las superficies dadas paramétricamente esto conduce a las superficies continuamente derivables. La definición de las últimas está basada sobre el concepto de aplicaciones equivalentes respecto de las transformaciones continuamente derivables.

Demos a conocer este concepto. Por función continuamente derivable en la clausura de cierta región se entenderá, como antes (véase el p. 39.3), aquella que tiene derivadas continuas en la propia región las cuales pueden ser prolongadas a la frontera de la misma.

Una aplicación de cierta región cerrada se denomina *continuamente derivable*, si toda función coordenada que determinada dicha aplicación (véase el p. 41.4), es continuamente derivable en la región cerrada que se considera. Las funciones prolongadas en estos casos se denotan mediante los mismos símbolos que se usan para designar funciones prolongables de partida.

Si cierta aplicación $u_1 = \varphi(u, v)$, $v_1 = \psi(u, v)$ es continuamente derivable en la clausura $\bar{D}$ de la región D, entonces, en virtud del convenio adoptado, esto es testimonio de que, en particular, el jacobiano $\dfrac{\partial(\varphi, \psi)}{\partial(u, v)}$ de esta aplicación es continuamente prolongable de la región D a su clausura $\bar{D}$ y la prolongación del jacobiano, que se designa por el mismo símbolo $\dfrac{\partial(\varphi, \psi)}{\partial(u, v)}$, se llamará también jacobiano.

Ante todo es necesario enunciar qué se entenderá por aplicaciones equivalentes continuamente derivables; con este fin introduzcamos el concepto de aplicaciones regulares.

Definición 8. *Una aplicación homeomorfa F de la clausura $\bar{D}$ de una región plana D sobre la clausura $\bar{D}_1$ de otra región plana D_1 que hace pasar los puntos interiores a los interiores y los puntos de frontera a los de frontera, lleva el nombre de aplicación regular de la región cerrada $\bar{D}$ sobre otra región cerrada $\bar{D}_1$, si tanto la propia aplicación F, como la aplicación F^{-1}, inversa de F, son continuamente derivables en las regiones cerradas $\bar{D}$ y $\bar{D}_1$, respectivamente.*

Observemos que toda aplicación regular F de una región cerrada $\bar{D}$ tiene en todos los puntos de D un jacobiano distinto de cero. En efecto, de acuerdo con la definición 8, en la aplicación F la imagen de cada punto interior es un punto interior. Por cuanto en estos puntos las aplicaciones directa y, correspondientemente, inversa son continuamente derivables, sus jacobianos no pueden reducirse a cero, pues el producto de éstos es igual a uno (véase el p. 41.7).

De aquí se deduce que el jacobiano de una aplicación regular F tampoco es nulo en la región cerrada $\overline{D}$. Efectivamente, en virtud de que los jacobianos de las aplicaciones, tanto directa, como inversa, son continuamente prolongables a las clausuras respectivas $\overline{D}$ y $\overline{F(D)}$ de las regiones D y $F(D)$, el producto de dichos jacobianos es también igual a uno para todos los puntos de la región cerrada $\overline{D}$.

Ya nos encontramos con las aplicaciones regulares de regiones planas cerradas del tipo especial, por ejemplo, en el p. 46.1.

Definición 9. *Supongamos que f y f_1 son las aplicaciones continuas de las clausuras $\overline{D}$ y $\overline{D}_1$ de las regiones planas D y D_1 en el espacio R^3 y sean dichas aplicaciones continuamente derivables en las regiones cerradas $\overline{D}$ y $\overline{D}_1$. Las aplicaciones f y f_1 se llaman equivalentes respecto de las aplicaciones continuamente derivables, si existe tal aplicación regular F de la región cerrada $\overline{D}$ sobre otra región cerrada $\overline{D}_1$ que para todo punto $M \in \overline{D}$ se cumple la condición* (50.1).

Ahora podemos definir la superficie continuamente derivable.

Definición 10. *Cualquier conjunto de aplicaciones $r(u, v)$ de las regiones planas cerradas $\overline{D}$ en el espacio tridimensional R^3, siendo dichas aplicaciones continuamente derivables y equivalentes respecto de las transformaciones continuamente derivables, lleva el nombre de superficie continuamente derivable S, dada en forma paramétrica, y cada una de las apicaciones mencionadas $r(u, v)$ se denomina representación de esta superficie*

$$S = \{r(u, v);\ (u, v) \in \overline{D}\}.$$

Subrayemos que si la superficie $S = \{r(u, v);\ (u, v) \in D\}$ es continuamente derivable, esto significa, en particular, que cualquier su representación vectorial $r = r(u, v)$, $(u, v) \in \overline{D}$, tiene derivadas parciales r_u y r_v*) que son continuas en la región D y pueden ser continuamente prolongadas a la frontera de ésta. Como, de acuerdo con el convenio adoptado, las funciones prolongadas se designan mediante los mismos símbolos que las prolongables **), se puede considerar que las funciones r_u y r_v son continuas en la región cerrada $\overline{D}$.

De modo semejante podemos definir también otras clases de las superficies dadas paramétricamente, por ejemplo, las superficies dadas en forma paramétrica que

---

*) Los conceptos de tal género como son, por ejemplo, continuidad, límite, derivabilidad se extienden de modo natural a las funciones vectoriales de varias variables. Así, la función $r(u, v)$, definida en la región G, se llama *continua en el punto* $(u_0, v_0) \in G$, si $\lim_{(u,v)\to(u_0,v_0)} r(u, v) = r(u_0, v_0)$. La derivada $\dfrac{\partial r}{\partial u}(u_0, v_0)$ se determina por la igualdad

$$\frac{\partial r(u_0, v_0)}{\partial u} = \left.\frac{dr(u, v_0)}{du}\right|_{u=u_0}$$

**) Con más precisión, este convenio se ha adoptado (véase el p. 39.3) para las funciones escalares y, por lo tanto, para las coordenadas de las funciones vectoriales, razón por la cual es natural adoptarlo para las mismas funciones vectoriales.

son dos veces continuamente derivables o, en general, n veces continuamente derivables, como también el concepto de su punto, portador y su parte.

Al resumir, podemos decir en definitiva que se llama *superficie de tal o cual clase, dada en forma paramétrica, cierta totalidad de aplicaciones* $r(u, v)$, $(u, v) \in \bar{D}$, *equivalentes entre sí en determinado sentido, llamadas representaciones de la citada totalidad.*

El concepto de equivalencia se determina en dependencia de la clase que se elige.

Definición 11. *Las transformaciones de los parámetros las cuales realizan el paso de una representación de la superficie a otra, equivalente a la primera, se denominan admisibles.*

De este modo, si $r(u, v)$, $(u, v) \in \bar{D}$, y $\rho(u_1, v_1)$, $(u_1, v_1) \in \bar{D}_1$, son dos representaciones de una misma superficie de cierta clase, dada paramétricamente, y la aplicación

$$u_1 = \varphi(u, v),$$
$$v_1 = \psi(u, v)$$

de la región cerrada $\bar{D}$ sobre otra región cerrada $\bar{D}_1$ es la transformación admisible de los parámetros, entonces para todos los puntos $(u, v) \in \bar{D}$ se verifica la correlación (véase (50.1))

$$r(u, v) = \rho[\varphi(u, v), \psi(u, v)].$$

Una superficie dada en forma paramétrica se determina unívocamente, para la clase prefijada de transformaciones admisibles, por cualquiera de sus representaciones, razón por la cual con el fin de definir tal superficie basta prefijar sólo una de sus representaciones.

50.3. SUPERFICIES DADAS IMPLÍCITAMENTE

Demos a conocer un enfoque más al concepto de superficie. Si $F(x, y, z)$ es una función continua en cierta región tridimensional, la totalidad de los puntos (x, y, z) tales que

$$F(x, y, z) = 0, \tag{50.6}$$

se denomina *superficie dada implícitamente*. Sin detenernos detalladamente en el análisis del enfoque mencionado al concepto de superficie, diremos sólo que en el caso cuando la función F satisfaga en cierto punto (x_0, y_0, z_0) las condiciones del teorema sobre funciones implícitas (véase el p. 41.1), una parte de la superficie (50.6) en cierto entorno del punto citado (es decir, la intersección de este entorno con la superficie dada) admite una representación explícita y podemos decir que en tales condiciones una superficie, dada implícitamente, se reduce localmente a la superficie que se determina mediante una representación explícita (véase el p. 50.1). Precisamente sólo con este tipo de superficies dadas implícitamente se chocará en lo que sigue, por lo cual no nos detendremos especialmente en la explicación de tales o cuales conceptos para las superficies dadas en forma implícita.

A título de ejemplo más sencillo de una superficie dada implícitamente indiquemos la ecuación

$$x^2 + y^2 + z^2 = 1.$$

Los puntos cuyas coordenadas satisfacen esta ecuación forman la superficie de una bola de radio unidad con centro en el origen de coordenadas.

En adelante se estudiarán, principalmente, sólo las superficies continuas, dadas mediante representaciones paramétricas y provistas, en el caso general, de puntos múltiples. Se denominarán, como ya se ha indicado, simplemente "superficies"; si el concepto de superficie debe entenderse en otro sentido, esto se especificará especialmente.

50.4. PLANO TANGENTE Y NORMAL A LA SUPERFICIE

Sea

$$S = [r(u, v); (u, v) \in \bar{D}] \tag{50.7}$$

una superficie continuamente derivable. Examinemos cierta representación vectorial suya $r = r(u, v)$, $(u, v) \in \bar{D}$. Al igual que cualquiera de sus representaciones vectoriales, es una función vectorial continuamente derivable en la región plana cerrada $\bar{D}$.

Convengamos en considerar, para simplificar, que la intersección de cada recta $u = u_0$ o $v = v_0$ con la región cerrada $\bar{D}$ consta de un segmento (que, quizás, se degenera en un punto) o es vacía. Supongamos, por ejemplo, que la intersección de $\bar{D}$ con la recta $v = v_0$ es no vacía, entonces

$$r = r(u, v_0), \quad (u, v_0) \in \bar{D}$$

(v_0 es fijo) es la representación de cierta curva continuamente derivable que se llama *línea coordenada* (*u-línea*). El vector

$$r_u = \frac{\partial r}{\partial u} = (x_u, y_u, z_u)$$

será su vector tangente. Análogamente se determinan otras líneas coordenadas (*v-líneas*) con ayuda de la representación

$$r = r(u_0, v), \quad (u_0, v) \in \bar{D}$$

(u_0 es fijo) y los vectores, tangentes a ellas,

$$r_v = \frac{\partial r}{\partial v} = (x_v, y_v, z_v).$$

Definición 12. *Un punto $r(u, v)$ de la superficie* (50.7), *para el cual los vectores r_u y r_v no son colineales* (*linealmente independientes*) *se denomina no singular para la representación dada de esta superficie. En el caso contrario, es decir, cuando los vectores r_u y r_v son colineales en el punto dado, se denomina punto singular de la superficie para la representación dada de ésta.*

Si un punto de la superficie no es singular, en él, en particular, $r_u \neq 0$, $r_v \neq 0$. Evidentemente, un punto de la superficie es no singular, para la representación dada de la misma, cuando, y sólo cuando, en este punto $r_u \times r_v \neq 0$.

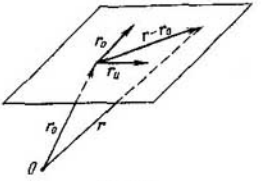


Fig. 213

Ejercicio 1. Demuéstrese que si $r(u_0, v_0)$ es un punto interior no singular de la superficie S, para la representación dada $r(u, v)$, $(u, v) \in \bar{D}$, es decir, para este punto $(u_0, v_0) \in D$ y $r_u \times r_v \neq 0$, entonces cierta parte de la superficie S, para la cual el punto $r(u_0, v_0)$ es también interior, cuenta con una representación explícita respecto de uno de los ejes de coordenadas.

Consideraremos una curva en la superficie (50.7). Supongamos que esta curva viene dada por las funciones continuamente derivables

$$u = u(t), \quad v = v(t), \quad (u(t), \quad v(t)) \in D, \quad a \leqslant t \leqslant b,$$

es decir, por la representación

$$r = r[u(t), v(t)], \quad (u(t), v(t)) \in D, \quad a \leqslant t \leqslant b, \tag{50.8}$$

con la particularidad de que $u'^2(t) + v'^2(t) > 0$ en $[a, b]$.

Al derivar la igualdad (50.8), obtendremos

$$dr = r_u du + r_v dv, \tag{50.9}$$

donde $du = u'(t)dt$, $dv = v'(t)dt$. Si el punto de la superficie, en el que se considera la igualdad (50.9), no es singular, el vector dr es tangente a la curva (50.8). La igualdad (50.9) muestra que en el punto dado $r(u_0, v_0)$ de la superficie (50.7) la tangente a cualquier curva (50.8) en dicha superficie que pasa por el punto $r(u_0, v_0)$ se dispone en el plano de los vectores $r_u(u_0, v_0)$ y $r_v(u_0, v_0)$.

Definición 13. *Un plano que contiene el punto $r(u_0, v_0)$ de la superficie* (50.7) *y en el que se disponen todas las tangentes a las curvas* (50.8) *que pasan por el punto citado, se denomina plano tangente a la superficie en el punto dado, llamado punto de tangencia.*

Ejercicio 2. Demuéstrese que para todo vector ν, dispuesto en el plano tangente a la superficie S en un punto no singular, existe en la superficie S una curva que pasa por el punto dado, y para la cual el vector ν es una tangente.

Si el punto dado de la superficie (50.7) no es singular, en él siempre existe un plano tangente y, además, es único: a saber, en virtud de (50.9), de éste último sirve el plano que pasa por $r(u_0, v_0)$ siendo paralelo a los vectores $r_u(u_0, v_0)$ y $r_v(u_0, v_0)$. De aquí resulta fácil escribir su ecuación en la forma vectorial. Al designar con r_0 el radio vector del punto de tangencia y con r, el radio vector corriente de los puntos en el plano tangente, obtendremos (fig. 213):

$$(r - r_0)r_u r_v = 0$$

(en el primer miembro de la igualdad figura un producto mixto de los vectores mencionados).

Si $r = (x, y, z)$, $r_0 = (x_0, y_0, z_0)$, $r_u = (x_u, y_u, z_u)$ y $r_v = (x_v, y_v, z_v)$, la ecuación del plano tangente en la forma coordenada se escribirá del modo siguiente:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} = 0.$$

En el caso de que la superficie sea dada explícitamente

$$z = f(x, y), \quad (x, y) \in \bar{D}, \tag{50.10}$$

tendremos $u = x$, $v = y$, por lo cual

$$\begin{gathered} x_u = 1, \quad y_u = 0, \quad z_u = f_x, \\ x_v = 0, \quad y_v = 1, \quad z_v = f_y; \end{gathered} \tag{50.11}$$

por consiguiente, la ecuación del plano tangente en este caso tendrá por expresión

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = 0,$$

de donde

$$z - z_0 = (x - x_0)f_x + (y - y_0)f_y, \tag{50.12}$$

donde con f_x y f_y están designadas, para abreviar, las derivadas parciales $f_x(x, y)$ y $f_y(x, y)$ en el punto (x_0, y_0).

De esta fórmula se deduce que las dos definiciones del plano tangente para una superficie representada explícitamente en la forma (50.10), dadas en el presente punto y en el p. 20.5, son equivalentes. En efecto, ambas definiciones conducen a una misma ecuación (50.12).

Definición 14. *Una recta que pasa por el punto de tangencia de una superficie con el plano tangente y que es perpendicular a este plano, se denomina recta normal a la superficie en el punto citado.*

Su ecuación, en el caso general, tiene por expresión en un punto no singular de la superficie

$$\frac{x - x_0}{\begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix}} = \frac{y - y_0}{\begin{vmatrix} z_u & x_u \\ z_v & x_v \end{vmatrix}} = \frac{z - z_0}{\begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}}.$$

En el caso de la representación explícita (50.10) estas ecuaciones asumen la forma

$$\frac{x - x_0}{f_x} = \frac{y - y_0}{f_y} = -(z - z_0). \tag{50.13}$$

Definición 15. *Todo vector no nulo, colineal con una recta normal, que pasa por el punto dado de una superficie, se denomina normal a dicha superficie en el punto citado.*

Como ejemplo de normal en un punto no singular de una superficie sirve el producto escalar

$$n = r_u \times r_v,$$

calculado en el punto mencionado.

De acuerdo con la definición aducida, en todo punto no singular (para la representación dada) $r(u, v)$ de la superficie en consideración existe, siendo fijados los valores de los parámetros u y v, una recta normal y dicha recta es, además, única. Se debe tener en cuenta que si el punto P de un espacio es un punto múltiple de la superficie, es decir, si existen por lo menos dos pares de parámetros (para la representación dada) (u_1, v_1) y (u_2, v_2) tales que $P = r(u_1, v_2) = r(u_2, v_2)$, entonces puede, naturalmente, ocurrir que a estos pares de parámetros les correspondan diferentes rectas normales, a consecuencia de lo cual en el punto citado P la recta normal no será única.

Para una superficie dada implícitamente por la ecuación

$$F(x, y, z) = 0,$$

donde $F(x, y, z)$ es una función continuamente derivable en el entorno del punto (x_0, y_0, z_0), $F(x_0, y_0, z_0) = 0$, y en este punto $F_x^2 + F_y^2 + F_z^2 > 0$, la ecuación del plano tangente en el punto (x_0, y_0, z_0) tiene por expresión

$$(x - x_0)F_x + (y - y_0)F_y + (z - z_0)F_z = 0,$$

donde F_x, F_y y F_z representan los valores de las derivadas parciales correspondientes tomadas en el punto (x_0, y_0, z_0).

Al recordar que un vector de coordenadas F_x, F_y, F_z, es decir, el vector $\nabla F = (F_x, F_y, F_z)$ lleva el nombre de *gradiente de la función* F (véase el p. 20.6), vemos que el gradiente de la función en el punto dado de la superficie $F(x, y, z) = 0$ es perpendicular al plano tangente en este punto, es decir, es colineal con la recta normal.

Por ello la ecuación de la recta normal a la superficie tiene por expresión

$$\frac{x - x_0}{F_x} = \frac{y - y_0}{F_y} = \frac{z - z_0}{F_z}.$$

Todas estas fórmulas se infieren inmediatamente de (50.12) y (50.13). En efecto, si, por ejemplo, $F_z \neq 0$ y $z = f(x, y)$ es una función definida por la ecuación $F = 0$ en el entorno del punto (x_0, y_0, z_0), entonces basta notar que $f_x = -\frac{F_x}{F_z}$, $f_y = -\frac{F_y}{F_z}$ (véase el p. 41.1).

Si la función dada $F(x, y, z)$ es continuamente derivable en la región G, para cualquier punto de una superficie dada implícitamente por la ecuación $F(x, y, z) = c$ (c es una constante), obtendremos la ecuación del plano tangente y de la recta normal de la misma forma que en el caso $F = 0$, siempre que en dicho punto $F_x^2 + F_y^2 + F_z^2 > 0$. Un conjunto de los puntos $(x, y, z) \in G$, para los cuales $F = c$, se denomina, como ya sabemos, superficie de nivel de la función F (véase el p. 19.1).

De este modo, el gradiente $\nabla F = (F_x, F_y, F_z)$ en el punto (x_0, y_0, z_0) de la superficie de nivel $F(x, y, z) = c$ está orientado a lo largo de la recta normal a dicha superficie en el punto (x_0, y_0, z_0). En otras palabras, el gradiente de la función es ortogonal a la superficie de nivel (es decir, es perpendicular al plano tangente a la superficie de nivel en el punto que se considera).

Hemos demostrado la existencia del plano tangente en un punto no singular perteneciente a una superficie continuamente derivable cuando esté fijada la representación de la misma. Surge una cuestión: ¿qué será, si pasamos a otra representación de esta superficie? Ante todo, ¿seguirá siendo no singular un punto no singular, y, un punto singular, singular? Resulta que sí.

Demostrémoslo. Sean $r(u, v)$, $(u, v) \in \bar{D}$, y $\rho(u_1, v_1)$, $(u_1, v_1) \in \bar{D}_1$, dos representaciones de una misma superficie continuamente derivable. Por cuanto el paso de cualquier representación de una superficie continuamente derivable a otra representación suya se realiza por medio de la aplicación regular, existe tal aplicación regular

$$\begin{aligned} u_1 &= \varphi(u, v), \\ v_1 &= \psi(u, v) \end{aligned} \tag{50.14}$$

de la región cerrada $\bar{D}$ sobre otra región cerrada $\bar{D}_1$ que para todos los puntos $(u, v) \in \bar{D}$ resulta ser válida la igualdad

$$r(u, v) = \rho[\varphi(u, v), \psi(u, v)]. \tag{50.15}$$

En este caso según lo demostrado más arriba, el jacobiano de la aplicación (50.14) nunca se reduce a cero en la región cerrada $\bar{D}$:

$$\frac{\partial(\varphi, \psi)}{\partial(u, v)} = \begin{vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{vmatrix} \neq 0, \quad (u, v) \in \bar{D}.$$

Al derivar la identidad (50.15), obtendremos

$$\begin{aligned} r_u &= \varphi_u \rho_{u_1} + \psi_u \rho_{v_1}, \\ r_v &= \varphi_v \rho_{u_1} + \psi_v \rho_{v_1}. \end{aligned} \tag{50.16}$$

Por consiguiente, un par de vectores ρ_{u_1}, ρ_{v_1} se transforma en otro par de vectores r_u, r_v con ayuda de una matriz regular

$$\begin{Vmatrix} \varphi_u & \psi_v \\ \varphi_u & \psi_v \end{Vmatrix}.$$

Por ello, para el punto dado (u, v) los vectores r_u, r_v serán linealmente independien-

tes, si, y sólo si, son linealmente independientes los vectores ρ_{u_1}, ρ_{v_1} en el punto (u_1, v_1) obtenido del punto (u, v) con ayuda de la aplicación (50.14), con la particularidad de que en el caso de su independencia lineal el plano de los vectores r_u y r_v coincide con el de los vectores ρ_{u_1} y ρ_{v_1}.

Así pues, *un punto (para la representación dada) no singular (singular) de una superficie continuamente derivable será también no singular (singular) para otra representación de la superficie citada, y un plano, tangente a la superficie en un punto no singular para una representación de la superficie, será tangente también para otra representación de la misma.*

Definición 16. *Una superficie continuamente derivable privada de puntos singulares, se llama superficie suave.*

De acuerdo con lo demostrado más arriba, para comprobar que la superficie dada es suave, basta convencerse de que la misma cuenta con una representación continuamente derivable y para esta última no hay puntos singulares.

Conviene fijar la atención en que las funciones vectoriales r_u y r_v de la superficie suave $S = [r(u, v), (u, v) \in \overline{D}]$ no sólo son continuas en la clausura de la región D, sino, de conformidad con la definición, no son colineales en esta clausura $\overline{D}$. En otras palabras, para la superficie suave (50.7) en todo punto de la región cerrada $\overline{D}$ se cumple la desigualdad

$$r_u \times r_v \neq 0.$$

OBSERVACIÓN. De las fórmulas (50.16) proviene que

$$r_u \times r_v = (\varphi_u \rho_{u_1} + \psi_u \rho_{v_1}) \times (\varphi_v \rho_{u_1} + \psi_v \rho_{v_1}) =$$
$$= \varphi_u \psi_v (\rho_{u_1} \times \rho_{v_1}) + \psi_u \varphi_v (\rho_{v_1} \times \rho_{u_1}) =$$
$$\frac{\partial(\varphi, \psi)}{\partial(u, v)} (\rho_{u_1} \times \rho_{v_1}).$$

Por cuanto, para las transformaciones admisibles de los parámetros (50.14), el jacobiano $\dfrac{\partial(\varphi, \psi)}{\partial(u, v)}$ nunca se reduce a cero en $\overline{D}$, de la fórmula obtenida se deduce que los productos vectoriales $r_u \times r_v$ y $\rho_{u_1} \times \rho_{v_1}$ en el punto dado de la superficie pueden reducirse a cero, si sólo se reducen simultáneamente. Pero, se ha notado que la condición necesaria y suficiente para que el punto dado de la superficie, para la representación dada de la superficie $r(u, v)$, sea no singular es que el producto vectorial $r_u \times r_v$ en dicho punto sea distinto de cero. De este modo hemos demostrado una vez más que un punto no singular (singular) de la superficie será el mismo tanto para una representación de la superficie, como para la otra.

Ejercicios. 3. Fórmense las ecuaciones de un plano tangente y de una normal a la superficie $x = 2u - v$, $y = u^2 + v^2$, $z = u^3 - v^3$ en el punto $M(3; 5; 7)$.

4. Trácese a la superficie $xyz = 1$ un plano tangente que sea paralelo al plano $x + y + z - a = 0$ (a = const).

5. Demuéstrese que todos los planos tangentes de la superficie $z = xf\left(\dfrac{y}{x}\right)$ (f es una función derivable arbitraria) pasan por el origen de coordenadas.

6. Demuéstrese que todos los planos tangentes, trazados a la superficie $x = u\cos v$, $y = u \operatorname{sen} v$, $z = au + f(v)$ ($a = \text{const}$, f es una función derivable arbitraria) en todo punto de su línea coordenada $v = c$ ($c = \text{const}$), pasan a través de la recta fija.

50.5. PRIMERA FORMA CUADRÁTICA DE UNA SUPERFICIE

Fijemos una representación $r = r(u, v)$, $(u, v) \in \bar{D}$, de la superficie suave dada y analicemos un plano tangente a la superficie en alguno de sus puntos. Según hemos visto, los vectores r_u y r_v forman en dicho plano una base. Los vectores, dispuestos en el plano tangente, los designaremos por el símbolo dr, y sus coordenadas respecto de la base r_u y r_v, por du y dv *). De este modo,

$$dr = r_u du + r_v dv.$$

Hallemos el cuadrado de la longitud de un vector, dispuesto en el plano tangente, expresándolo en términos de las coordenadas de la base natural r_u y r_v (en el álgebra lineal esta expresión se denomina, comúnmente, *forma métrica principal* del espacio en consideración, en el caso dado, de un plano)

$$|dr|^2 = (r_u du + r_v dv)^2 = r_u^2 du^2 + 2r_u r_v du dv + r_v^2 dv^2.$$

Introduzcamos las designaciones

$$E = r_u^2, \quad F = r_u r_v, \quad G = r_v^2; \tag{50.17}$$

entonces

$$|dr|^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2. \tag{50.18}$$

Definición 17. *Una forma cuadrática* $E\,du^2 + 2F\,du\,dv + G\,dv^2$ *se denomina primera forma cuadrática de la superficie* **).

Veamos cómo varía pasando a la otra representación de la superficie (véanse las fórmulas (50.14)). Como se sabe (véase (50.16)) en este caso las bases en el plano que se considera se transforman con ayuda de una matriz

$$\begin{Vmatrix} \varphi_u & \psi_u \\ \varphi_v & \psi_v \end{Vmatrix}.$$

Por consiguiente, las coordenadas de los vectores se transforman con ayuda de una matriz transpuesta, es decir, de la matriz de Jacobi

$$J = \begin{Vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{Vmatrix}.$$

---

*) Esta designación es bien natural, pues si un vector en el plano tangente es tangente a cierta curva (50.8) en la superficie, entonces, al elegir adecuadamente el parámetro, el vector dr será una diferencial del vector (50.8) y, por lo tanto, para él se verificará la igualdad (50.9).

**) El hecho de que la forma cuadrática en consideración se denomina primera se debe a que existen también otras formas cuadráticas relacionadas con la superficie. El estudio de éstas sale de los márgenes de este libro.

Al designar mediante A la matriz de la primera forma cuadrática (50.18) para la representación de la superficie $r = r(u, v)$, y mediante A_1, para la representación $\rho = \rho(u_1, v_1)$, es decir,

$$A = \begin{Vmatrix} E & F \\ F & G \end{Vmatrix}, \quad E = r_u^2, \quad F = r_u r_u r_v, \quad G = r_v^2,$$

$$A_1 = \begin{Vmatrix} E_1 & F_1 \\ F_1 & G_1 \end{Vmatrix}, \quad E_1 = \rho_{u_1}^2, \quad F_1 = \rho_{u_1}\rho_{v_1}, \quad G_1 = \rho_{v_1}^2,$$

entonces, según se sabe por el curso del álgebra lineal, para la primera forma cuadrática de la superficie, al igual que, en general, para cualquier forma cuadrática,

$$A = J^* A_1 J,$$

donde con J^* está designada la matriz transpuesta respecto de la matriz de Jacobi J.

De aquí, para los determinantes correspondientes

$$\begin{vmatrix} E & F \\ F & G \end{vmatrix} = \begin{vmatrix} E_1 & F_1 \\ F_1 & G_1 \end{vmatrix} \begin{vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{vmatrix}^2,$$

o bien

$$EG - F^2 = (E_1 G_1 - F_1^2) \left| \frac{\partial(u_1, v_1)}{\partial(u, v)} \right|^2. \tag{50.19}$$

Ha de notarse que en virtud de la propia definición, la primera forma cuadrática es definida positiva (en efecto, si $du^2 + dv^2 > 0$, es decir, si $dr \neq 0$, entonces $|dr|^2 > 0$) y por esta razón su discriminante es positivo: $EG - F^2 > 0$. Estando ausentes los puntos singulares, se verifican las desigualdades $r_u \neq 0$, $r_v \neq 0$, por lo cual, de la definición de coeficientes E y G (50.17), proviene inmediatamente que $E > 0$ y $G > 0$.

Si se conoce la primera forma cuadrática de una superficie, podemos resolver (sin disponer de la ecuación de la superficie y sin conocer las formas de la misma), toda una serie de problemas concernientes a ella, por ejemplo, hallar las longitudes de las curvas dispuestas en la superficie, como también los ángulos entre ellas, calcular el área de las partes de la superficie. La variedad de todas las propiedades de la superficie que pueden establecerse, partiendo sólo de la primera forma cuadrática, se denomina *geometría interior* de la superficie. Pasemos precisamente a la consideración de los problemas similares.

Ejercicios. 7. ¿Cuál de las siguientes formas cuadráticas puede servir de primera forma cuadrática de cierta superficie: a) $du^2 + 3dudv + dv^2$; b) $du^2 + 6dudv + 9dv^2$; c) $du^2 - 6dudv + 13dv^2$; d) $du^2 + 2dudv - dv^2$.

8. Hállese la primera forma cuadrática de un *helicoide* (superficie helicoidal) $x = u\cos v$, $y = u \operatorname{sen} v$, $z = av + f(u)$ ($a = \text{const}$, f es una función derivable arbitraria).

9. Demuéstrese que la primera forma cuadrática de una superficie de revolución es reducible a la forma $du^2 + G(u)dv^2$.

50.6. CURVAS EN UNA SUPERFICIE. CÁLCULO DE SUS LONGITUDES Y DE ÁNGULOS ENTRE ELLAS

Examinemos una curva continuamente derivable (50.8) dispuesta en la superficie dada (50.7). Supongamos que las longitudes de los arcos $s = s(t)$ en la superficie se calculan en el sentido de crecimiento del parámetro, es decir, que $\frac{ds}{dt} > 0$. Según se sabe (véase el p. 16.5), $\frac{ds}{dt} = \left|\frac{dr}{dt}\right|$, de donde $ds = |dr|$, por consiguiente, véase (50.18),

$$ds^2 = |dr|^2 = dr^2 = E\,du^2 + 2F du dv + G dv^2,$$

por eso

$$\frac{ds}{dt} = \sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\frac{du}{dt}\frac{dv}{dt} + G\left(\frac{dv}{dt}\right)^2}.$$

De este modo, para la longitud L de la curva (50.8) obtenemos la fórmula

$$L = \int_a^b \sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\frac{du}{dt}\frac{dv}{dt} + G\left(\frac{dv}{dt}\right)^2}\,dt.$$

Pasemos ahora al cálculo de los ángulos entre las curvas en una superficie.

Definición 18. *Si dos curvas se cortan en cierto punto, el ángulo entre ellas en el punto citado será aquel que está formado por sus tangentes en el punto mencionado (si es que existen).*

Supongamos que dos curvas suaves, dispuestas en la superficie en consideración, se cortan en cierto punto. Designemos las diferenciales de sus representaciones en dicho punto mediante dr y δr, respectivamente, y los coeficientes de las descomposiciones según los vectores r_u y r_v, mediante du, dv y δu, δv, respectivamente; en este caso

$$dr = r_u du + r_v dv,$$

$$\delta r = r_u \delta u + r_v \delta v.$$

Por eso, si φ es el ángulo buscado entre las curvas, es decir, el formado por los vectores dr y δr, entonces

$$\cos\varphi = \frac{dr\,\delta r}{|dr|\,|\delta r|} = \frac{E du\delta u + F(du\delta v + dv\delta u) + G dv\delta v}{\sqrt{E du^2 + 2F du dv + G dv^2}\,\sqrt{E\delta u^2 + 2F\delta u\delta v + G\delta v^2}}.$$

Ejercicios. 10. Demuéstrese que para que las u-líneas y v-líneas coordenadas en una superficie sean ortogonales, es necesario y suficiente que en todo punto de la superficie se verifique la igualdad $F = 0$.

11. Hállese el ángulo entre las curvas $v = 2u$, $v = -2u$ en una superficie con la primera forma cuadrática $du^2 + dv^2$.

50.7. AREA DE UNA SUPERFICIE

Supongamos que la representación continuamente derivable $r(u, v)$ de la superficie suave en consideración S se ha definido en la clausura $\bar{D}$ de la región cuadrable D. Examinemos la partición T_k del plano de variables u y v en cuadrados de cierto rango k. Por cuanto del carácter cuadrable de la región proviene el carácter acotado de ésta, la región cerrada $\bar{D}$ resultará cubierta por un número finito de cuadrados de rango k. Numeremos de uno u otro modo todas las intersecciones no vacías de estos cuadrados con la región cerrada $\bar{D}$ y designémoslas mediante X_i, $i = 1, 2, \ldots, i_0$. Entonces

$$\tau = \{X_i : X_i = Q \cap \bar{D} \neq \phi, Q \in T_k\}$$

forma la partición de la región cerrada $\bar{D}$ (véase la definición de la partición en el p. 44.3).

Examinemos los conjuntos X_i que representan en sí cuadrados cerrados completos dispuestos en la región D (siendo suficientemente menuda la finura de la partición τ, tales conjuntos no vacíos X_i siempre existen; ¿por qué?). Denotemos con $\tau(\partial D)$ la totalidad de todos los conjuntos mencionados X_i (compárese con el p. 44.4).

Elijamos un cuadrado cualquiera $X_i \in \tau(\partial D)$ (fig. 214). Sea h la longitud de su lado y sea P_i uno de sus vértices. Entonces, al pasar del vértice P_i a los vértices vecinos, el radio vector $r(u, v)$ adquirirá (con exactitud de hasta los infinitésimos de orden superior que h) los incrementos iguales en valor absoluto a los números $|r_u h|$ y $|r_v h|$, respectivamente, pues

$$r(u + h, v) - r(u, v) = r_u h + o(h),$$
$$r(u, v + h) - r(u, v) = r_v h + o(h).$$

Al determinar el área de la superficie, sustituyamos las imágenes de los cuadrados $X_i \in \tau(\partial D)$ por los paralelogramos rectilíneos construidos a partir de los vectores $r_u h$ y $r_v h$ (fig. 215). Hallemos el área de tal paralelogramo. Al designarla con $\Delta\sigma_i$, obtendremos

$$\Delta\sigma_i = |r_u h \times r_v h|_{P_i} = |r_u \times r_v|_{P_i} h^2 = |r_u \times r_v|_{P_i} \mu X_i.$$

Las funciones r_u y r_v son continuas en la región cerrada cuadrable $\bar{D}$, por lo cual

$$\lim_{\delta_\tau \to 0} \sum_{X_i \in \tau(\partial D)} \Delta\sigma_i = \iint_{\bar{D}} |r_u \times r_v| du dv, \tag{50.20}$$

donde δ_τ significa, como siempre, la finura de la partición τ. Evidentemente, la condición de que la finura de la partición δ_τ tiende a cero es equivalente a que los rangos k de las particiones del plano en cuadrados, de las cuales hemos partido, tienden hacia el infinito.

Para demostrar la validez de la igualdad (50.20), basta observar que, siendo elegidos arbitrariamente los puntos $P_i \in X_i \in \tau$, $i = 1, 2, \ldots, i_0$, se verifica la igualdad

$$\lim_{\delta_\tau \to 0} \sum_{X_i \in \tau(\partial D)} \Delta\sigma_i = \lim_{\delta_\tau \to 0} \sum_{i=1}^{i_0} |r_u \times r_v|_{P_i} \mu X_i = \iint_{\bar{D}} |r_u \times r_v| du dv.$$

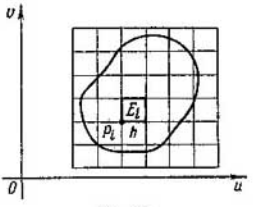


Fig. 214

Efectivamente, en primer lugar, el límite de las sumas integrales de una función integrable no depende de la elección, en el caso dado, de los puntos $P_i \in X_i \in \tau$, y, en segundo lugar, la omisión en las sumas integrales de los sumandos, correspondientes a los conjuntos $X_i \in \tau$ que no figuran en $\tau(\partial D)$, no influye, como se sabe (véase el p. 44.3), en la magnitud del límite de sumas integrales, en nuestro caso, en la magnitud del límite (50.20).

Definición 19. *El límite* (50.20) *lleva el nombre de área o medida* μS *de la superficie* S:

$$\mu S = \lim_{\sigma_\tau \to 0} \sum_{X_i \in \tau(\partial D)} \Delta\sigma_i.$$

Para calcular el área de la superficie, de (50.20) se obtiene inmediatamente la fórmula

$$\mu S = \iint_D |r_u \times r_v| \, du dv. \tag{50.21}$$

Escribámosla en otra forma, expresando el integrando en términos de los coeficientes de la primera forma cuadrática. Diremos, ante todo, que para cualesquiera vectores a y b son lícitas las fórmulas

$$|a \times b| = |a|\,|b| \operatorname{sen} \widehat{ab},$$
$$ab = |a|\,|b| \cos \widehat{ab},$$

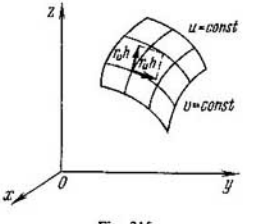


Fig. 215

donde $\widehat{ab}$ es el ángulo formado por los vectores a y b. Elevemos al cuadrado y sumemos estas fórmulas:

$$|a \times b|^2 + |ab|^2 = a^2b^2.$$

De aquí se deduce que

$$|r_u \times r_v|^2 = r_u^2 r_v^2 - (r_u r_v)^2 = EG - F^2, \tag{50.22}$$

por lo cual la fórmula (50.21) puede escribirse también así:

$$\mu S = \iint_D \sqrt{EG - F^2}\, du\, dv. \tag{50.23}$$

A veces, para abreviar la notación, la expresión $\sqrt{EG - F^2}\, du\, dv$ se designa por el símbolo dS:

$$dS = \sqrt{EG - F^2}\, du\, dv; \tag{50.24}$$

y se llama *elemento del área*. Aplicando esta designación, podemos escribir la fórmula (50.23) en la forma

$$\mu S = \iint_D dS.$$

Mostremos que la magnitud del área de una superficie no depende de la elección de su representación (en este caso se consideran sólo las representaciones dadas en las regiones cerradas cuadrables). Pasemos a otra representación $\rho = \rho(u_1, v_1)$ de la superficie continuamente derivable dada, la cual viene dada en la clausura $\bar{D}_1$ de la región cuadrable D_1 y, por consiguiente, para la cual la transformación (50.14) de los parámetros u, v en los u_1, v_1 es una aplicación regular de D sobre $\bar{D}_1$.

En el nuevo sistema de coordenadas consideraremos la integral

$$\mu S = \iint_{D_1} \sqrt{E_1 G_1 - F_1^2}\, du_1 dv_1.$$

Para compararla con la integral (50.23), realicemos el cambio de variables (50.14), lo que es posible, puesto que todas las premisas del teorema 2° del p. 46.2 están en este caso cumplidas. Al utilizar (50.19), obtendremos

$$\mu S_1 = \iint_{D_1} \sqrt{E_1 G_1 - F_1^2}\, du_1 dv_1 =$$

$$= \iint_D \sqrt{E_1 G_1 - F_1^2} \left| \frac{\partial(\varphi, \psi)}{\partial(u, v)} \right| du\, dv = \iint_D \sqrt{EG - F^2}\, du\, dv = \mu S.$$

De este modo, efectivamente, la magnitud del área de una superficie no depende de cómo se elige su representación.

Hallemos la expresión para el área de una superficie que tiene representación explícita $z = f(x, y)$, $(x, y) \in \bar{D}$. En este caso $u = x$, $v = y$, $r = (x, y, f(x, y))$ y, por lo tanto (véanse las fórmulas (50.11)),

$$r_u = (1, 0, f_x), \quad r_v = (0, 1, f_y),$$

$$E = r_u^2 = 1 + f_x^2, \quad F = r_u r_v = f_x f_y, \tag{50.25}$$

$$G = r_v^2 = 1 + f_y^2, EG - F^2 = (1 + f_x^2)(1 + f_y^2) - \\ - f_x^2 f_y^2 = 1 + f_x^2 + f_y^2; \mu S = \iint_D \sqrt{1 + f_x^2 + f_y^2}\, dxdy. \quad (50.25)$$

Ejercicios. 12. Demuéstrese que el área de una superficie de revolución, definida en el p. 32.4, coincide con el área de esta superficie, definida en el punto presente.

13. Hállense el perímetro y los ángulos interiores de un triángulo curvilíneo dispuesto en la superficie con la primera forma cuadrática $du^2 + (u^2 + a^2)dv^2$ y limitado por los arcos de las curvas $u = \frac{1}{2}av^2$, $u = -\frac{1}{2}av^2$, $v = 1$ ($a = \text{const} > 0$).

14. Hállese el área de un cuadrilátero curvilíneo dispuesto en el *helicoide* $x = u \cos v$, $y = u \operatorname{sen} v$, $z = au$ ($a = \text{const}$) y limitado por los arcos de las curvas $u = 0$, $u = a$, $v = 0$, $v = 1$.

15. En una superficie con la primera forma cuadrática $du^2 + (u^2 + a^2)dv^2$ está dispuesto un triángulo curvilíneo limitado por los arcos de las curvas $u = av$, $u = -av$, $v = 1$. Hállese el área de este triángulo.

50.8. ORIENTACIÓN DE LA SUPERFICIE SUAVE

En este párrafo se supondrá que en un espacio siempre se elige el sistema directo de coordenadas rectangulares cartesianas. Esto significa lo siguiente. Sean i, j y k los versores de los ejes coordenados. Mirando desde el extremo del vector k en la dirección del plano xOy, concluimos que con el fin de hacer coincidir el vector i con el j se debe girar el primero al ángulo $\frac{\pi}{2}$ en el sentido contrahorario. En este caso suele decirse también que la terna ordenada de vectores i, j, k está concordada según la "regla de sacacorchos". Analíticamente esto significa que en el espacio de puntos (x, y, z) se consideran sólo tales bases ordenadas e_1, e_2, e_3 que se obtienen de la base ordenada $i = (1; 0; 0)$, $j = (0; 1; 0)$, $k = (0; 0; 1)$ con ayuda de las matrices que tienen determinante positivo (con más precisión, igual a + 1). De esta forma, si

$$e_m = c_{m1}i + c_{m2}j + c_{m3}k, \quad m = 1, 2, 3,$$

es una base que define el sistema directo de coordenadas, entonces

$$\begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix} = +1.$$

Todas las definiciones y nociones que se introducen más abajo en este párrafo y que están relacionadas con las coordenadas vienen dadas con arreglo a los sistemas directos de coordenadas.

Sea S una superficie suave (véase la definición 16). Cualquier su representación vectorial $r = r(u, v)$, $(u, v) \in \overline{D}$, es continuamente derivable y $r_u \times r_v \neq 0$ en la región cerrada $\overline{D}$.

Por consiguiente, en todo punto de la superficie S está definido el vector unidad normal

$$\nu = \frac{r_u \times r_v}{|r_u \times r_v|}, \quad (50.26)$$

que es una función continua en $\bar{D}$. Esta circunstancia se expresa brevemente diciendo que en la superficie S existe una normal unitaria continua.

Definición 20. *Toda normal unitaria continua* $\nu = \nu(u, v)$, $(u, v) \in \bar{D}$, *de una superficie suave* $S = \{r(u, v); (u, v) \in \bar{D}\}$ *lleva el nombre de orientación de la superficie* S.

Evidentemente, si el vector ν es una orientación de la superficie S, $-\nu$ será también orientación de la misma superficie y es fácil mostrar que otras orientaciones no existen.

Ejercicio 16. Demuéstrese que una superficie puede tener sólo dos orientaciones.

Una de las dos orientaciones, ν ó $-\nu$ (elegida arbitrariamente) se llama *positiva* y la otra, *negativa*.

De este modo, el concepto de orientación positiva o negativa en este sentido no se determina univocamente por la propia superficie, sino depende de la elección de su representación. Las orientaciones positiva y negativa de una superficie se llaman orientaciones *opuestas* de la misma.

En lo que sigue, para concretar, *como* orientación positiva, para una superficie suave dada mediante la representación vectorial fija $r = r(u, v)$, $(u, v) \in \bar{D}$, se tomará siempre el vector (50.26).

Subrayemos que la continuidad de la normal ν se considera respecto de las variables u, v, y no respecto de las variables espaciales x, y, z. Si la superficie cuenta con puntos múltiples, puede ocurrir que en el punto de un espacio, portador de distintos puntos de la superficie, pueden haber varias normales distintas.

Cuando se requiere que, en la transformación regular de los parámetros u, v, la superficie conserve la orientación, se debe exigir adicionalmente que el jacobiano de tal transformación sea positivo. En efecto, para la transformación de los parámetros

$$u_1 = \varphi(u, v),$$
$$v_1 = \psi(u, v)$$

de las fórmulas (50.16), según lo hemos visto (véase la observación al final del p. 50.4), se deduce que

$$r_u \times r_v = \frac{\partial(\varphi, \psi)}{\partial(u, v)} (\rho_{u_1} \times \rho_{v_1})$$

y, por consiguiente, si el jacobiano $\dfrac{\partial(\varphi, \psi)}{\partial(u, v)}$ es positivo, los vectores $r_u \times r_v$ y $\rho_{u_1} \times \rho_{v_1}$ están dirigidos hacia un mismo lado, y si es negativo, hacia los lados opuestos.

De este modo, para las superficies con la orientación elegida, como transformaciones admisibles se considerarán aquellas transformaciones continuamente derivables cuyo jacobiano es positivo.

La superficie S con una orientación positiva se designará mediante S^+, y con una orientación negativa, mediante S^-.

Subrayemos que cualquier superficie suave dada paramétricamente es siempre orientable, es decir, siempre tiene su orientación.

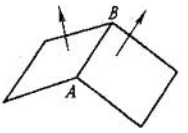


Fig. 216

Definición 21. *Una superficie que tiene fija una de sus orientaciones se denomina orientada.*

La definición de orientación, aducida anteriormente, no puede ser extendida, por supuesto, a las superficies no suaves. Como ejemplo de superficie no derivable en un punto, en la que ya no puede elegirse una normal continua, sirve el cono

$$z = \sqrt{x^2 + y^2}, \quad x^2 + y^2 \leqslant a^2. \tag{50.27}$$

En este caso la representación vectorial tiene por expresión:

$$r(x, y) = (x, y, \sqrt{x^2 + y^2}),$$

por consiguiente

$$r_x = \left(1; 0; \frac{x}{\sqrt{x^2 + y^2}}\right), \quad r_y = \left(0; 1; \frac{y}{\sqrt{x^2 + y^2}}\right);$$

$$r_x \times r_y = \left(-\frac{x}{\sqrt{x^2 + y^2}}; -\frac{y}{\sqrt{x^2 + y^2}}; 1\right); \quad |r_x \times r_y| = \sqrt{2}.$$

Por cuanto los límites $\lim\limits_{(x,y)\to(0,0)} \frac{x}{\sqrt{x^2 + y^2}}$ y $\lim\limits_{(x,y)\to(0,0)} \frac{y}{\sqrt{x^2 + y^2}}$ no existen (¿por qué?), entonces la normal unitaria

$$\nu = \frac{r_x \times r_y}{|r_x \times r_y|} = \left(-\frac{x}{\sqrt{2(x^2 + y^2)}}; -\frac{y}{\sqrt{2(x^2 + y^2)}}; \frac{1}{\sqrt{2}}\right)$$

tampoco tiene límite para $(x, y) \to (0; 0)$. Por ello, en el cono (50.27) no se puede elegir una normal continua en $\overline{D} = \{(x, y) : x^2 + y^2 \leqslant a^2\}$.

Como ejemplo de superficie no suave S, en la que existe toda una línea a lo largo de la cual las normales, cualquier que sea su elección, sufren discontinuidad, sirve una parte del ángulo diedro expuesta en la fig. 216. La línea mencionada en esta superficie es el segmento AB.

50.9. PEGAMIENTO DE LAS SUPERFICIES

La definición, aducida más arriba, de la superficie continua dada en la forma paramétrica no abarca todo lo que interviene intuitivamente en el concepto de superficie. Así, por ejemplo, podemos mostrar que la superficie de una bola no es portadora de ninguna superficie continua dada paramétricamente sin puntos múltiples. Por

otra parte, sería una complicación no justificada considerar que la superficie de una bola está provista de puntos múltiples. Existen varios procedimientos para vencer este inconveniente. Elegimos un procedimiento basado en el pegamiento de un número finito de superficies. El pegamiento de las superficies surge de un modo natural en la resolución de los más sencillos problemas. Por ejemplo, es natural considerar la superficie lateral de un cilindro como resultado de pegar los lados opuestos de un rectángulo; la superficie total de un cilindro, como resultado de pegar su superficie lateral y dos bases; la superficie de un cono, como resultado de pegar la superficie lateral de éste con su base, etc.

Pasemos a las definiciones precisas. Diremos que el borde (véase el p. 50.1) de la superficie $S = \{r = r(u, v); (u, v) \in \overline{D}\}$ es una curva, si la frontera ∂D de la región D es también una curva (con más precisión, portadora de una curva):

$$\partial D = \{u(t), v(t); a \leqslant t \leqslant b\}.$$

En este caso el borde ∂S de la superficie S puede considerarse como una curva:

$$\partial S = \{r(u(t), v(t)); a \leqslant t \leqslant b\}.$$

Definamos la operación de pegado de las superficies para aquellas superficies cuyos bordes son curvas.

Sean dadas las superficies $S_i = \{r_i(u_i, v_i); (u_i, v_i) \in \overline{D}_i\}$ cuyos bordes ∂S_i son unas curvas, es decir, lo son las fronteras ∂D_i de las regiones D_i:

$$u_i = u_i(t_i), \quad v_i = v_i(t_i), \quad a_i \leqslant t_i \leqslant b_i, \quad i = 1, 2, \ldots, m.$$

Entonces los bordes ∂S_i de las superficies serán las curvas

$$\gamma_i = \{r_i(u_i(t_i), v_i(t_i)); a_i \leqslant t_i \leqslant b_i\}.$$

Supongamos que para algunos pares (i, j), $i, j = 1, 2, \ldots, m$, $i \neq j$, se han dado un número finito de los segmentos $[a_{ij}^k, b_{ij}^k] \subset [a_i, b_i]$, $a_{ij}^k \leqslant b_{ij}^k$, y de los segmentos $[a_{ji}^k, b_{ji}^k] \subset [a_j, b_j]$, $a_{ji}^k \leqslant b_{ji}^k$, $k = 1, 2, \ldots, n_{ij} = n_{ji}$, con la particularidad de que tanto los segmentos $[a_{ij}^k, b_{ij}^k]$, como también los $[a_{ji}^k, b_{ji}^k]$ no tienen dos a dos puntos interiores comunes ni tampoco homeomorfismos $\varphi_{ij}^k: [a_{ij}^k, b_{ij}^k] \to [a_{ji}^k, b_{ji}^k]$, llamados *homeomorfismos de pegamiento*. Además, en este caso, para cualquier $t_i \in [a_{ij}^k, b_{ij}^k]$ tiene lugar el "pegado"

$$r_i(u_i(t_i), v_i(t_i)) = r_j(u_j(\varphi_{ij}^k(t_i)), \quad v_j(\varphi_{ij}^k(t_i))). \tag{50.28}$$

Designemos mediante γ_{ij}^k una curva cuya representación es

$$r_i(u_i(t_i), \quad v_i(t_i)), \quad t \in [a_{ij}^k, b_{ij}^k].$$

Las curvas γ_{ij}^k se llaman *curvas de pegamiento* o *curvas a lo largo de las cuales se realiza el pegamiento*.

Es evidente que, en virtud de (50.28), la aplicación

$$r = r_j(u_j(t_j), v_j(t_j)), \quad t_j \in [a_{ji}^k, b_{ji}^k],$$

es también la representación de la curva γ_{ij}^k, pues los homeomorfismos φ_{ij}^k representan en sí una transformación admisible del parámetro para la curva γ_{ij}^k.

Supondremos, además, que para $j' \neq j$ los segmentos

$$[a_{ij}^k, b_{ij}^k] \quad \text{y} \quad [a_{ij'}^l, b_{ij'}^l], \quad k = 1, 2, \ldots, n_{ij}, \quad l = 1, 2, \ldots, n_{ij'}$$

no tienen puntos interiores comunes y, por consiguiente, todo extremo del segmento $[a_{ij}^k, b_{ij}^k]$ puede pertenecer, además, a lo sumo a un segmento $[a_{ij'}^l, b_{ij'}^l]$. Esta condición significa que toda curva de pegamiento γ_{ij}^k es una parte sólo de dos curvas, γ_i y γ_j, que forman los bordes de las superficies S_i y S_j.

Las superficies S_i y S_j se llaman adyacentes, si se pegan por lo menos a lo largo de una curva γ_{ij}^k. El sistema de homeomorfismos de pegamiento φ_{ij}^k se denomina *conexo*, si para cualesquiera superficies S_p y S_q del sistema en consideración existen en él tales superficies $S_{i_1}, S_{i_2}, \ldots, S_{i_r}$, que $S_{i_1} = S_p$, $S_{i_r} = S_q$ y cada superficie S_{i_ν} es adyacente respecto de $S_{i_{\nu+1}}$, es decir, se ha pegado con ésta a lo largo de una o varias curvas con ayuda de los homeomorfismos de pegamiento correspondientes $\varphi_{i_\nu i_{\nu+1}}$, $\nu = 1, 2, \ldots, r-1$.

Definición 22. *Un sistema de superficies $S_1, S_2, \ldots, S_m$ con el sistema conexo de homeomorfismos de pegamiento φ_{ij}^k se llama superficie pegada de las superficies $S_1, \ldots, S_m$ a lo largo de las curvas γ_{ij}^k y se designa mediante $S = \{S_i\}$.*

Esta definición, a pesar de ser formalmente engorrosa, tiene, evidentemente, un significado geométrico muy simple. Hablando metafóricamente, la superficie pegada $S = \{S_i\}$ representa en sí las superficies $S_1, \ldots, S_m$, ciertos pares de las cuales S_i, S_j tienen identificados (pegados) los puntos que se disponen en las curvas γ_{ij}^k y se aplican uno en el otro en los homeomorfismos φ_{ij}^k, en lo que precisamente consiste la condición de pegamiento (50.28). Indudablemente, como ya se ha indicado, se supone, además, que de cada superficie S_i se puede pasar, realizados un número finito de pasos, a otra superficie cualquiera S_j, cada vez pasando de una cierta superficie a otra adyacente.

Si $S = \{S_i\}$ es una superficie pegada, la totalidad de todos los arcos, que representan tales partes de las curvas ∂S_i que ningunos puntos de las partes indicadas, a excepción, quizás, de los extremales, se pegan con los puntos de otras curvas ∂S_j, se denomina borde ∂S de la superficie pegada S.

Se puede mostrar que reuniendo de modo adecuado las partes citadas de las curvas ∂S_i, pertenecientes al borde ∂S de la superficie $S = \{S_i\}$, podemos obtener un número finito de curvas cerradas (contornos). En otras palabras, el borde de una superficie pegada consta de un número finito de contornos cerrados.

Como ejemplo de pegamiento de las superficies puede servir el pegamiento, para obtener una esfera $x^2 + y^2 + z^2 = 1$, de dos semiesferas $z = \sqrt{1 - x^2 - y^2}$ y $z = -\sqrt{1 - x^2 - y^2}$, $x^2 + y^2 \leqslant 1$, a lo largo de su borde, es decir, a lo largo de la circunferencia $x^2 + y^2 = 1$, $z = 0$. Al definir la ecuación de esta circunferencia en la forma paramétrica

$$x = \cos t, \quad y = \operatorname{sen} t, \quad z = 0, \; 0 \leqslant t \leqslant 2\pi,$$

a título de homeomorfismo de pegamiento φ: $[0, 2\pi] \to [0, 2\pi]$ podemos tomar la aplicación idéntica del segmento $[0, 2\pi]$ sobre sí mismo.

Con ayuda del pegamiento de las superficies suaves se puede dar la noción de superficie suave a trozos.

Definición 23. *Una superficie* $S = \{S_i\}$, *obtenida pegando las superficies suaves* $S_1, \ldots, S_m$, *se llama superficie suave a trozos.*

Como ejemplos de superficies suaves a trozos pueden servir la superficie de un cilindro circular, la de un paralelepípedo. Entre tanto, un cono recto circular (50.27) no puede ser dividido en un número finito de partes suaves pegadas, razón por la cual no es una superficie suave a trozos en el sentido de la definición 23. La operación de pegamiento de las superficies puede generalizarse de un modo tal que, conservándose intacta la definición de superficies suaves a trozos para tal operación generalizada de pegamiento, en la clase de las superficies suaves a trozos ya se incluirán también las superficies cónicas. No nos detendremos en esto dejando a cargo del propio lector de hacerlo, en caso de necesidad.

50.10. SUPERFICIES ORIENTABLES Y NO ORIENTABLES

Nuestra tarea inmediata consiste en dar la definición de orientación para las superficies, pegadas de unas superficies que se dan en la forma paramétrica.

La definición de la orientación eligiendo una normal unitaria continua en la superficie resulta en el caso dado incómoda, incluso cuando los puntos múltiples estén ausentes y la continuidad de la normal se entienda como si dependiese continuamente de los puntos del espacio (y no de los parámetros de las superficies que se pegan). Esto se debe a la perturbación eventual de la suavidad de la superficie en aquellas curvas a lo largo de las cuales se realiza el pegamiento.

Por ejemplo, una parte de la superficie del ángulo diedro expuesta en la fig. 216, puede considerarse como resultado de pegar dos rectángulos iguales. Al tender por distintas caras hacia el mismo punto en la arista de dicho ángulo, obtendremos distintos límites de las normales unitarias correspondientes. Más abajo se dará tal definición de una superficie orientable, en términos de la cual la superficie citada será orientable.

Observemos que incluso cuando el "procedimiento suave" se use para pegar las superficies (es decir, cuando para toda curva, a lo largo de la cual se realiza el pegamiento, podemos elegir, en cualquiera de sus puntos, una normal unitaria de modo tal que fuera ésta el límite para las normales unitarias, elégidas de modo adecuado en el entorno del punto citado, de dos superficies que se pegan) las superficies pegadas pueden adquirir nuevas peculiaridades cualitativas: a diferencia de las superficies dadas en la forma paramétrica, en el caso dado no siempre en toda la superficie puede elegirse una normal unitaria continua. Como ejemplo de tal superficie sirve la así llamada cinta o banda de Möbius *), la cual puede obtenerse de una tira rectangular de papel $ABCD$ que se tuerce una vez alrededor de su eje de simetría MN, paralela a los lados BC y AD, después de lo cual se pegan las aristas AB y CD (fig. 217). Verdad es que, empleado tal procedimiento, la cinta de Möbius se obtiene como resultado de pegar la superficie con sí misma. No obstante, no es difícil obtenerla también pegando, de conformidad con la definición 22, dos rectángulos $ABEF$ y $FECD$ (véase la fig. 217).

*) A. F. Möbius (1790—1868), matemático y astrónomo alemán.

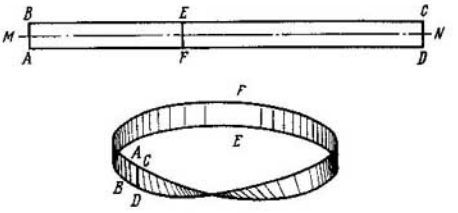


Fig. 217

Una de las peculiaridades características de la cinta de Möbius consiste en que ella tiene sólo una "cara": es imposible, como, por ejemplo, en el caso de la superficie lateral de un cilindro, pintarla de rojo un lado y de azúl, otro lado. Además, en la cinta de Möbius no puede elegirse una normal unitaria que fuera una función continua de un punto del espacio.

Todos los razonamientos aducidos hacen naturales las tentativas de ofrecer tal definición de orientación de una superficie, para la cual las superficies del tipo, por ejemplo, de la superficie de un paralelepípedo resultarían orientadas, y las superficies del tipo de la cinta de Möbius, no orientadas.

Prestemos atención a que la cinta de Möbius puede ser un pordador de una superficie suave, dada paramétricamente, con puntos múltiples y esta superficie, al igual que cualquier otra superficie suave dada paramétricamente, será orientada. Por supuesto, esto no tiene nada que ver con el carácter no orientado de la propia cinta de Möbius.

50.11. SEGUNDO ENFOQUE DEL CONCEPTO DE ORIENTACIÓN DE UNA SUPERFICIE

Pasemos ahora a la descripción de otro enfoque del concepto de orientación que se basa en el pegamiento de las superficies cuyos bordes son unas curvas. Sea $S = \{r = r(u, v); (u, v) \in \overline{D}\}$ una superficie suave cuyo borde está constituido por una curva. La orientación positiva de la curva $\partial D = \{u(t), v(t); a \leqslant b \leqslant b\}$ (es decir, la orientación contrahoraria en el plano u, v con el sistema directo de coordenadas) engendra, en virtud de la aplicación $r(u(t), v(t))$, $a \leqslant t \leqslant b$, una orientación bien determinada del borde ∂S de la superficie S. Esta orientación del borde ∂S de la superficie S se llama concordada con la orientación

$$\nu = \frac{r_u \times r_v}{|r_u \times r_v|}$$

(véase la definición 20) de la superficie S.

El carácter natural de dicha definición puede aclararse del modo siguiente. Consideremos una superficie dada explícitamente $S: z = f(x, y)$, $(x, y) \in \overline{D}$. Para ella

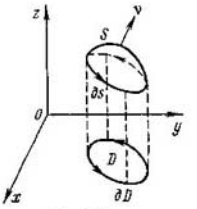


Fig. 218

(véase (50.10), (50.11) y (50.26))

$$\nu = \left(-\frac{f_x}{\sqrt{1+f_x^2+f_y^2}}, -\frac{f_y}{\sqrt{1+f_x^2+f_y^2}}, \frac{1}{\sqrt{1+f_x^2+f_y^2}} \right).$$

Por consiguiente, $\cos(\widehat{\nu, k}) = \dfrac{1}{\sqrt{1+f_x^2+f_y^2}} > 0$, es decir, el vector de la normal ν forma con el eje Oz un ángulo agudo y, por eso, está concordado con la orientación positiva del borde ∂S de la superficie S según la regla de sacacorchos: la orientación del contorno ∂S corresponde al sentido de rotación de la manija del sacacorchos y la dirección de la normal ν, al movimiento del propio sacacorchos (véase la fig. 218).

Evidentemente, si la orientación ν de la superficie suave en consideración S está concordada con la de su borde ∂S, entonces la orientación $-\nu$ queda concordada con la orientación opuesta de la curva ∂S. De este modo, la definición de la orientación ν de una superficie suave es equivalente a la definición de la orientación de la curva ∂S que es el borde de la superficie. Por eso el borde orientado ∂S de la superficie suave S lo llamaremos, al igual que la normal unitaria continua ν, *orientación de la superficie* S.

Para una superficie no suave, dada paramétricamente, cuyo borde está constituido por un contorno, su orientación puede considerarse como la definición de partida para la orientación de la propia superficie. Sean S_1 y S_2 dos superficies suaves cuyos bordes son unas curvas y supongamos que estas dos superficies están pegadas (en el sentido de la definición 22) a lo largo de las curvas $\gamma_1, \ldots, \gamma_m$, que constituyen las partes de los bordes de las superficies S_1 y S_2. Las orientaciones ∂S_1 y ∂S_2 de las superficies S_1 y S_2 se denominan *concordadas*, siempre que cada una de ellas engendra en las curvas $\gamma_1, \ldots, \gamma_m$ que se pegan las orientaciones opuestas.

Definición 24. *Una superficie* S, *pegada de las superficies* $S_1, \ldots, S_m$, *se denomina orientable, si existen tales orientaciones* $\partial S_1, \ldots, \partial S_m$ *de los contornos de las superficies* $S_1, \ldots, S_m$ *que para cualesquiera dos superficies adyacentes* S_i *y* S_j *sus orientaciones* ∂S_i *y* ∂S_j *están concordadas.*

La totalidad de tales orientaciones, si es que existe, se llama *orientación de la superficie* S.

Si la totalidad mencionada de orientaciones ∂S_i no existe, la superficie S se llama *no orientable*.

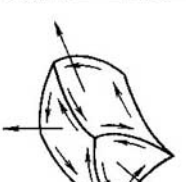


Fig. 219

Si $\partial S_1, \ldots, \partial S_m$ es la orientación de la superficie $S = \{S_i\}$, la totalidad de orientaciones opuestas es también una orientación de la superficie S, llamada orientación *opuesta respecto de la dada*.

Se puede mostrar que si la superficie S es orientable, ella no tiene ningunas orientaciones a excepción de las dos orientaciones mencionadas. Una de las dos orientaciones existentes (no importa cuál sea) se llama, comúnmente, positiva, y la otra, negativa.

Por analogía, con el caso examinado antes en el p. 50.8, una superficie orientable que tiene fija una de sus orientaciones se denomina *orientada*. Además, aquella de las superficies orientadas, cuya orientación se ha llamado positiva, se designa con S^+, y la superficie de orientación opuesta, con S^-.

El borde de una superficie pegada orientada $S = \{S_i\}$, como todo borde de la superficie pegada, se compone, de acuerdo con lo dicho anteriormente, de un número finito de contornos cerrados. Cada uno de estos contornos, a su vez, representa una reunión del número finito de curvas, cada una de las cuales forma parte de uno de los contornos ∂S_i, a saber, una parte tal que todos los puntos suyos, a excepción, quizás, de los extremos, no se pegan con los puntos de los otros bordes ∂S_j. Por esta razón, la orientación concordada dada de la superficie orientable pegada $S = \{S_i\}$ engendra determinadas orientaciones (es decir, los órdenes de los puntos) en las curvas mencionadas. Se puede mostrar que dichas orientaciones, tomadas juntas, constituyen las orientaciones de todos los contornos que integran el borde ∂S de la superficie pegada S. La totalidad de estas orientaciones de los contornos que componen el borde ∂S de la superficie S lleva el nombre de *orientación* de dicho *borde* engendrada por la orientación dada de la superficie S, o, que es lo mismo, concordada con ella.

Prestemos atención en que en la definición 24 de la orientación de una superficie ni siquiera se suponía la derivabilidad de las superficies a pegar $S_1, \ldots, S_m$.

Si la superficie S está pegada de las superficies suaves $S_1, \ldots, S_m$, entonces, para prefijar su orientación, se pueden elegir en cada superficie $S_1, \ldots, S_m$ las normales unitarias continuas de una manera tal que las orientaciones ∂S_i de los bordes de las superficies S_i, concordadas con dichas normales, sean concordadas entre sí en el sentido de la definición 24, es decir, que representen la orientación de la superficie S (véase la fig. 219).

Fig. 220

Para enterarse, con tal modo de dar la orientación, de si coinciden o no dos orientaciones, es suficiente recurrir tan sólo a un punto arbitrario; si las normales coinciden en dicho punto, coinciden en todo punto, siempre que, naturalmente, existen; si en el punto mencionado las normales no coinciden, es decir, son opuestas, serán opuestas en todo punto, pues, como lo hemos indicado más arriba, existen sólo dos orientaciones de la superficie dada.

Sin embargo, en el caso de una superficie suave a trozos ya no podemos introducir la noción de orientación positiva, haciendo uso de las representaciones dadas de las superficies suaves a pegar y tomando en éstas normales unitarias según la fórmula (50.26), puesto que estas orientaciones pueden ser no concordadas. Por ello, cuando se trata de las superficies suaves a trozos, se debe especificar cada vez qué precisamente se sobreentiende en el caso concreto por superficies orientadas S^+ y S^- de la superficie dada S.

Se puede mostrar que cualquier superficie suave a trozos que constituye la frontera de cierta región del espacio tridimensional es orientable. En este caso una de las orientaciones se compone de las normales unitarias dirigidas de la superficie al interior de la región (las así llamadas *normales interiores*), y la otra se compone de las normales unitarias dirigidas de la superficie al exterior de la región (las así llamadas *normales exteriores*). Como ejemplo de tal superficie interviene una esfera. En calidad de su orientación pueden tomarse, por ejemplo, las normales unitarias dirigidas a lo largo del radio desde un punto de la esfera hacia el centro (fig. 220).

Como ejemplo de superficie no orientable (en el sentido de la definición 24) sirve la cinta de Möbius.

A veces las superficies orientables suaves a trozos se llaman, además, *superficies bilaterales*: tienen dos "caras" correspondientes a dos elecciones de las normales unitarias que prefijan dos orientaciones suyas. Correspondientemente, las superficies no orientables se denominan *unilaterales*. El empleo de este término se ha explicado con el ejemplo de cinta de Möbius en el p. 50.10.

No nos detendremos aquí en los razonamientos matemáticos concernientes a la demostración de las afirmaciones enunciadas. Esto requeriría el empleo de métodos cuyo estudio sale de los márgenes de este curso. Las afirmaciones generales enunciadas más arriba sin demostración no se usan, en esencia, en la exposición ulterior. Cuando, en lo sucesivo, se trate de un caso concreto, se puede indicar siempre cuál precisamente orientación se considera.

Ejercicios. 17. Demuéstrese que el cilindro recto circular representa una superficie suave a trozos sin borde.

18. Sean dados el vector τ y la curva $\gamma = \{\rho(u),\ a \leqslant u \leqslant b\}$.

Una superficie, dada mediante la representación de la forma

$$r = r(u, v) \overset{\text{def}}{=} \rho(u) + v\tau, \quad a \leqslant u \leqslant b, \quad c \leqslant v \leqslant d,$$

lleva el nombre de *superficie cilíndrica* S con la generatriz γ y una directriz paralela al vector τ.

Demuéstrese que si la curva γ es suave a trozos, la superficie S será también suave a trozos.

§ 51. INTEGRALES DE SUPERFICIE

En este párrafo y en los que siguen se considerarán sólo las superficies definidas por las representaciones paramétricas y, además, aquellas que son suaves (véase la definición 16 en el § 50) y suaves a trozos (véase la definición 23, en el § 50).

51.1. DEFINICIONES Y PROPIEDADES DE LAS INTEGRALES DE SUPERFICIE

Sea dada una superficie suave S, con la particularidad de que

$$r = r(u, v) = \{x = x(u, v), \quad y = y(u, v), \quad z = z(u, v);\ (u, v) \in \overline{D}\} \tag{51.1}$$

es su representación, con más precisión, representación continuamente derivable sin puntos singulares, D es una región plana cuadrable y, como siempre, E, G y F son los coeficientes de la primera forma cuadrática de la superficie S. Supongamos, luego, que en el conjunto de puntos $r(u, v)$ de la superficie S está definida la función Φ, es decir, la función $\Phi(r(u, v)) = \Phi(x(u, v), y(u, v), z(u, v))$. A veces la función Φ se denotará también mediante $\Phi(x, y, z)$ (compárese con el p. 47.1).

Definición 1. *Una integral* $\iint\limits_S \Phi(x, y, z)dS$ *se determina por la igualdad* (véase (50.24))

$$\iint\limits_S \Phi(x, y, z)dS = \iint\limits_D \Phi(x(u, v), y(u, v), z(u, v))\sqrt{EG - F^2}\,du\,dv. \tag{51.2}$$

Ésta se llama *integral de superficie de primera especie*.

Con ciertas restricciones impuestas en la función Φ la integral (51.2) existe. Así, por ejemplo, existe para toda función Φ, continua en la superficie suave $S = \{r(u, v),\ (u, v) \in \overline{D}\}$, es decir, para la función $\Phi(r(u, v))$, continua en la región cerrada cuadrable $\overline{D}$. Efectivamente, en este caso, de acuerdo con la definición 1, la integral

$$\iint\limits_S \Phi(x, y, z)dS$$

se reduce a una integral de la función continua en $\overline{D}$, la cual, como se sabe (véase el p. 44.4), existe. Las condiciones más generales de existencia de la integral de superficie de primera especie pueden obtenerse de las condiciones correspondientes de

existencia de las integrales múltiples (véase el p. 44.4), aplicadas a la integral que figura en el segundo miembro de la igualdad (51.2).

Supongamos, para simplificar, que la función Φ es continua en la superficie suave S y sea $\rho = \rho(u_1, v_1) = (\varphi(u_1, v_1), \psi(u_1, v_1), \chi(u_1, v_1))$ otra representación de esta superficie; dicha representación viene dada en la clausura $\bar{D}_1$ de la región cuadrable $\bar{D}_1$ y para ella la transformación (50.14) de los parámetros u, v en u_1, v_1 es biunívoca y continuamente derivable en $\bar{D}$, y, además, su jacobiano en $\bar{D}$ es distinto de cero. Si E_1, F_1 y G_1 son coeficientes de la primera forma cuadrática, correspondientes a dicha representación, entonces

$$\iint_D \Phi(x(u, v), y(u, v), z(u, v)) \sqrt{EG - F^2}\,du\,dv = \iint_{D_1} \Phi(\varphi(u_1, v_1), \psi(u_1, v_1), \chi(u_1, v_1)) \sqrt{E_1 G_1 - F_1^2}\,du_1\,dv_1. \tag{51.3}$$

Para convencerse de esto, es suficiente realizar el cambio de variables (50.14) en la integral que figura en el segundo miembro de esta igualdad y hacer uso de la fórmula (50.19). De este modo, la integral de superficie de primera especie no depende de cómo se elige la representación de la superficie. Las integrales de superficie de primera especie se encuentran en varios problemas de las matemáticas y sus aplicaciones. Por ejemplo, el área de una superficie (véase el p. 50.7) se expresa con ayuda de una integral de superficie de primera especie: si la función $\Phi(x, y, z)$ es idénticamente igual a la unidad en la superficie S, la fórmula (51.2) se convierte en una que sirve para determinar el área μS de la superficie S (véase (50.23)):

$$\mu S = \iint_D \sqrt{XG - F^2}\,du\,dv = \iint_S dS.$$

Si $\Phi(x, y, z)$ es la densidad de cierta masa distribuida por la superficie S, entonces la integral (51.2) proporciona la magnitud de la masa de toda la superficie.

Sean ahora, como siempre, i, j y k los vectores unidad coordenados,

$$n = r_u \times r_v = \begin{vmatrix} i & j & k \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} = \frac{\partial(y, z)}{\partial(u, v)} i + \frac{\partial(z, x)}{\partial(u, v)} j + \frac{\partial(x, y)}{\partial(u, v)} k \tag{51.4}$$

y

$$\nu = n/|n|, \tag{51.5}$$

con la particularidad de que, con arreglo a nuestras suposiciones, la normal ν es continuamente prolongable a la frontera de la región D.

La superficie S en la que se ha elegido la normal unitaria ν se designara con S^+, y la misma superficie con la normal elegida $-\nu$, con S^- (evidentemente, ν y $-\nu$ son dos orientaciones de la superficie S). Subrayemos que S^+ y S^- se determinan por la propia superficie "con la exactitud hasta la orientación" y dependen de la elección de la representación de la superficie.

Definición 2. *Las integrales de superficie*

$$\iint_{S^+} \Phi(x, y, z)\,dx\,dy \quad y \quad \iint_{S^-} \Phi(x, y, z)\,dx\,dy, \tag{51.6}$$

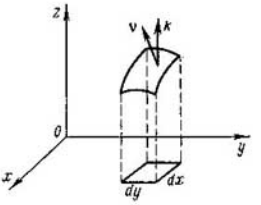


Fig. 221

llamadas integrales de superficie de segunda especie (para la representación dada de la superficie) se determinan según las fórmulas

$$\iint\limits_{S^+} \Phi(x,y,z)dxdy = \iint\limits_{S} \Phi(x,y,z)\cos(\widehat{\nu,k})dS,$$
$$\iint\limits_{S^-} \Phi(x,y,z)dxdy = \iint\limits_{S} \Phi(x,y,z)\cos(-\widehat{\nu,k})dS, \qquad (51.7)$$

donde (ν, k) *y* $(-\nu, k)$ *son los ángulos entre los vectores* ν, k *y entre* $-\nu, k$, *respectivamente*.

Como base de esta definición se ha tomado un razonamiento intuitivo de que el elemento del área dS de la superficie dada (véase (50.24)), multiplicado por el coseno del ángulo que este elemento "forma" con el plano xOy, es igual aproximadamente al elemento del área $dxdy$ de este plano (fig. 221), como si se tratase de las áreas de una figura plana y sus proyecciones.

Las integrales (51.6) se designarán mediante el símbolo común

$$\iint\limits_{S} \Phi(x,y,z)dxdy. \qquad (51.8)$$

Dado que $(\widehat{\nu,k}) + (-\widehat{\nu,k}) = \pi$, y, por lo tanto, $\cos(-\widehat{\nu,k}) = -\cos(\widehat{\nu,k})$ de (51.7) obtenemos

$$\iint\limits_{S^-} \Phi(x,y,z)dxdy = -\iint\limits_{S^+} \Phi(x,y,z)dxdy. \qquad (51.9)$$

Por analogía, con las integrales de superficie de primera especie, las integrales de superficie de segunda especie existen a ciencia cierta, si la función Φ es continua en la superficie S.

Por cuanto las integrales de superficie de primera especie no dependen de la representación de la superficie, las integrales de superficie de segunda especie (51.6) no dependen de cómo se elige la representación de la superficie orientada (en otras palabras, no dependen de cómo se elige la representación de una superficie que conserva su orientación dada), pero, por supuesto, las integrales (51.8) dependen en el caso general, para la superficie S y la función Φ dadas, de la elección de la normal

continua ν en la superficie, es decir, de la elección de la orientación de la superficie (véase 51.9).

Obtengamos las fórmulas que sean cómodas para calcular las integrales de superficie de segunda especie. Demos a conocer previamente que de (51.4), (51.5) y (50.22) proviene que

$$\cos(\widehat{\vec{\nu}, \vec{k}}) = \nu k = \frac{nk}{|n|} = \frac{(r_u \times r_v)k}{|r_u \times r_v|} = \frac{\dfrac{\partial(x, y)}{\partial(u, v)}}{|r_u \times r_v|} = \frac{\dfrac{\partial(x, y)}{\partial(u, v)}}{\sqrt{EG - F^2}},$$

razón por la cual $\sqrt{EG - F^2}\cos(\widehat{\vec{\nu}, \vec{k}}) = \dfrac{\partial(x, y)}{\partial(u, v)}$, y, por lo tanto,

$$\iint\limits_S \Phi(x, y, z)dxdy = \iint\limits_S \Phi(x, y, z)\cos(\widehat{\vec{\nu}, \vec{k}})dS =$$

$$= \iint\limits_D \Phi[x(u, v), y(u, v), z(u, v)]\cos(\widehat{\vec{\nu}, \vec{k}})\sqrt{EG - F^2}dudv =$$

$$= \iint\limits_D \Phi[x(u, v), y(u, v), z(u, v)]\frac{\partial(x, y)}{\partial(u, v)}dudv.$$

De este modo, omitiendo las designaciones de los argumentos de la función, tenemos

$$\iint\limits_{S_+} \Phi dxdy = \iint\limits_D \Phi \frac{\partial(x, y)}{\partial(u, v)}dudv \tag{51.10}$$

y, de conformidad con (51.9),

$$\iint\limits_{S_-} \Phi dxdy = -\iint\limits_D \Phi \frac{\partial(x, y)}{\partial(u, v)}dudv = \iint\limits_D \Phi \frac{\partial(y, x)}{\partial(u, v)}dudv. \tag{51.11}$$

A veces la integral $\iint\limits_{S_+} \Phi dxdy$ se designa como $\iint\limits_S \Phi dxdy$; en este caso la integral $\iint\limits_{S_-} \Phi dxdy$ se escribe en la forma $\iint\limits_S \Phi dydx$.

De este modo,

$$\iint\limits_S \Phi dxdy = \iint\limits_D \Phi \frac{\partial(x, y)}{\partial(u, v)}dudv,$$

$$\iint\limits_S \Phi dydx = \iint\limits_D \Phi \frac{\partial(y, x)}{\partial(u, v)}dudv.$$

Si la superficie S está dada explícitamente por la función continuamente derivable $z = f(x, y)$, $(x, y) \in \bar{D}$, entonces la fórmula (51.2) adquiere la forma (véase (50.25))

$$\iint_S \Phi(x, y, z)dS = \iint_D \Phi(x, y, f(x, y))\sqrt{1 + f_x^2 + f_y^2}dxdy,$$

mientras que las fórmulas (51.10) y (51.11) tendrán por expresión

$$\iint_{S^+} \Phi(x, y, z)dxdy = \iint_D \Phi(x, y, f(x, y))dxdy,$$

$$\iint_{S^-} \Phi(x, y, z)dxdy = -\iint_D \Phi(x, y, f(x, y))dxdy.$$

Aquí S^+ se llama "*cara superior de la superficie S*" (corresponde a la orientación positiva ν de la superficie S para la representación dada de ésta: $z = f(x, y)$), y S^-, "*cara inferior de la superficie S*" (corresponde a la orientación negativa $-\nu$). Estas denominaciones se deben a la circunstancia de que en el caso de definición explícita de la superficie

$$n = \begin{vmatrix} i & j & k \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = -f_x i + f_y j + k,$$

y, por consiguiente,

$$\nu = \left(-\frac{f_x}{\sqrt{1 + f_x^2 + f_y^2}}, -\frac{f_y}{\sqrt{1 + f_x^2 + f_y^2}}, \frac{1}{\sqrt{1 + f_x^2 + f_y^2}}\right),$$

y por ello, como ya se ha observado antes,

$$\cos(\widehat{\nu, k}) = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}} > 0.$$

De aquí se ve que el ángulo entre los vectores ν y k es agudo, es decir, el vector ν está dirigido "hacia arriba" de la superficie en consideración (véase fig. 221).

Por analogía con la definición (51.7) se determinan también otras integrales de superficie de segunda especie:

$$\begin{aligned}
\iint_{S^+} \Phi(x, y, z)dydz &= \iint_S \Phi(x, y, z)\cos(\widehat{\nu, i})dS, \\
\iint_{S^-} \Phi(x, y, z)dydz &= \iint_S \Phi(x, y, z)\cos(-\widehat{\nu, i})dS, \\
\iint_{S^+} \Phi(x, y, z)\,dzdx &= \iint_S \Phi(x, y, z)\cos(\widehat{\nu, j})dS, \\
\iint_{S^-} \Phi(x, y, z)\,dzdx &= \iint_S \Phi(x, y, z)\cos(-\widehat{\nu, j})dS.
\end{aligned} \tag{51.12}$$

Para estas integrales, por analogía con lo hecho anteriormente, obtendremos

$$\iint_{S^-} \Phi dydz = -\iint_{S^+} \Phi dydz,$$

$$\iint_{S^-} \Phi dzdx = -\iint_{S^+} \Phi dzdx,$$

$$\iint_{S^+} \Phi dydz = \iint_D \Phi \frac{\partial(y, z)}{\partial(u, v)} dudv, \tag{51.13}$$

$$\iint_{S^+} \Phi dzdx = \iint_D \Phi \frac{\partial(z, x)}{\partial(u, v)} dudv.$$

Los diversos problemas que conducen al concepto de integral de superficie de segunda especie serán considerados en el § 52.

51.2. INTEGRALES DE SUPERFICIE COMO LÍMITES DE LAS SUMAS INTEGRALES

Las integrales de superficie se pueden obtener también como límites de las sumas integrales correspondientes. Supongamos que S es una superficie suave, $r = r(u, v)$, $(u, v) \in \bar{D}$, la representación de dicha superficie y D, una región cuadrable. Supondremos, para simplificar, que en D existen particiones tan menudas como se quiera y los elementos de éstas son regiones cuadrables. Solamente las particiones de este género se considerarán en el presente punto. Tomemos cualquiera de las particiones mencionadas $\tau = \{D_i\}_{i=1}^{i=i_0}$ de la región D. Designemos mediante S_i, $i = 1, \ldots, i_0$, una superficie que se da por la representación $r = r(u, v)$, $(u, v) \in \bar{D}_i$. Es evidente que todas las S_i son también superficies suaves (el sistema $\tau_S = \{S_i\}_{i=1}^{i=i_0}$ lleva el nombre de *partición de la superficie* S). Supongamos que la función $\Phi(r(u, v)) = \Phi(x(u, v), y(u, v), z(u, v))$ es continua en $\bar{D}$ y $(u_i, v_i) \in \bar{D}_i$, $\Phi_i = \Phi(r(u_i, v_i))$. Denotemos con $\cos_i(\widehat{\nu, k})$ el coseno del ángulo formado por la normal ν y el versor k en el punto $r(u_i, v_i)$ de la superficie dada y pongamos

$$\sigma_\tau^{(1)} = \sum_{i=1}^{i_0} \Phi_i \mu S_i, \quad \sigma_\tau^{(2)} = \sum_{i=1}^{i_0} \Phi_i \cos_i(\widehat{\nu, k}) \mu S_i;$$

entonces

$$\lim_{\delta_\tau \to 0} \sigma_\tau^{(1)} = \iint_S \Phi(x, y, z) dS, \tag{51.14}$$

$$\lim_{\delta_\tau \to 0} \sigma_\tau^{(2)} = \iint_S \Phi(x, y, z)\, dxdy, \tag{51.15}$$

donde, como siempre, δ_τ es la finura de la partición τ. En efecto,

$$\iint_S \Phi(x, y, z) dS = \iint_D \Phi(r(u, v)) \sqrt{EG - F^2} dudv =$$

$$= \sum_{i=1}^{i_0} \iint_{D_i} \Phi(r(u, v)) \sqrt{EG - F^2} dudv;$$

por cuanto $\mu S_i = \iint\limits_{D_i} \sqrt{EG - F^2}\, du dv$, entonces

$$\sigma_\tau^{(1)} = \sum_{i=1}^{i_0} \Phi_i \iint\limits_{D_i} \sqrt{EG - F^2}\, du dv = \sum_{i=1}^{i_0} \iint\limits_{D_i} \Phi(r(u_i, v_i))\sqrt{EG - F^2}\, du dv.$$

Al designar ahora con $\omega(\delta; \Phi)$ el módulo de continuidad de la función Φ en la región cerrada $\bar{D}$, tendremos

$$\left| \iint\limits_S \Phi(x, y, z) dS - \sigma_\tau^{(1)} \right| \leqslant \sum_{i=1}^{i_0} \iint\limits_{D_i} |\Phi(r(u, v)) - \Phi(r(u_i, v_i))| \sqrt{EG - F^2}\, du dv \leqslant \omega(\delta_\tau, \Phi) \sum_{i=1}^{i_0} \mu S_i = \omega(\delta_\tau; \Phi)\mu S.$$

Al pasar en esta desigualdad al límite, para $\delta_\tau \to 0$, y observar que $\lim\limits_{\delta_\tau \to 0} \omega(\delta_\tau; \Phi) = 0$, obtendremos la fórmula (51.14).

Análogamente se demuestra la fórmula (5.15) (el producto $\Phi \cos(r, k)$ es continuo y, por ende, continuo uniformemente en $\bar{D}$). Los razonamientos similares son lícitos también para las integrales de segunda especie de otros tipos (15.12).

Ejercicio 1. Demuéstrese la fórmula (51.15).

51.3. INTEGRALES DE SUPERFICIE EXTENDIDAS A LAS SUPERFICIES SUAVES A TROZOS

Definamos las integrales de superficie extendidas a las superficies suaves a trozos.

Definición 3. *Supongamos que $S = \{S_i\}_{i=1}^{i=k}$ es una superficie suave a trozos (véase la definición 23 en el p. 50.9) y $\Phi(x, y, z)$, una función definida en el conjunto de puntos de la superficie S. Entonces, según la definición,*

$$\iint\limits_S \Phi dS = \sum_{i=1}^{k} \iint\limits_{S_i} \Phi dS_i. \tag{51.16}$$

Definición 4. *Si la superficie suave a trozos $S = \{S_i\}_{i=1}^{i=k}$ es orientable y $S^+ = \{S_i^+\}_{i=1}^{i=k}$ es una de las superficies orientadas que corresponde a S (véanse las designaciones en el p. 50.11), entonces, por definición,*

$$\iint\limits_{S^+} \Phi dx dy = \sum_{i=1}^{k} \iint\limits_{S_i^+} \Phi dx dy,$$

$$\iint_{S^+} \Phi dydz = \sum_{i=1}^{k} \iint_{S_i^+} \Phi dydz,$$

$$\iint_{S^+} \Phi dzdx = \sum_{i=1}^{k} \iint_{S_i^+} \Phi dzdx. \qquad (51.17)$$

Por supuesto, esta definición tiene sentido sólo en el caso en que las integrales en los segundos miembros de las igualdades existen. Para esto es necesario, ante todo, que las representaciones de las superficies S_i sean dadas en las regiones cuadrables.

Análogamente en este caso se definen también las integrales extendidas a la superficie $S^- = \{S_i^-\}_{i=1}^{i=k}$.

Nos hemos detenido sólo en aquellas propiedades de las integrales de superficie que están ligadas con el carácter específico de su definición y con la superficie a la que, como suele decirse, se extiende la integración. Es natural que por cuanto las integrales de superficie se reducen a las integrales múltiples ordinarias, para ellas son propias también las diferentes propiedades de las últimas (linealidad, teorema integral del valor medio, etc.).

OBSERVACIÓN. Las condiciones, enunciadas anteriormente (véanse las definiciones 10 en el p. 50.2 y 16 en el p. 50.4), que se imponen sobre las aplicaciones que realizan las transformaciones admisibles de los parámetros para las superficies suaves, resultan ser a menudo demasiado rígidas (compárese con la circunstancia similar para las curvas en el p. 47.3). Por ejemplo, adoptado tal procedimiento, las representaciones de la parte de una bola de radio unidad con centro en el origen de coordenadas, dispuesta en el primer octante:

$$z = \sqrt{1 - x^2 - y^2}, \text{ donde } x^2 + y^2 \leqslant 1, \quad x \geqslant 0, \quad y \geqslant 0$$

y

$$x = r\cos\psi\cos\varphi, \; y = r\cos\psi\,\operatorname{sen}\varphi, \quad z = r\operatorname{sen}\psi,$$

donde $0 \leqslant \varphi \leqslant \dfrac{\pi}{2}$, $0 \leqslant \psi \leqslant \dfrac{\pi}{2}$, no son equivalentes. Más aún, la primera representación no define, en el sentido indicado, una superficie continuamente derivable, por cuanto las derivadas parciales de la función $z = \sqrt{1 - x^2 - y^2}$ no están acotadas en la región $D = \{(x, y) : x^2 + y^2 < 1, x > 0, y > 0\}$ y no pueden ser continuamente prolongadas a su clausura $\bar{D}$. Resulta, pues, conveniente ampliar la definición de superficie continuamente derivable. Hagámoslo del modo siguiente.

Examinemos una totalidad de representaciones $r = r(u, v)$ $(u, v) \in \bar{D}$, continuas en $\bar{D}$ y continuamente derivables en D. Llamaremos transformaciones admisibles de los parámetros $u = \varphi(u_1, v_1), v = \psi(u_1, v_1), (u_1, v_1) \in \bar{D}_1$, a toda aplicación continua y biunívoca de la clausura $\bar{D}_1$ de la región plana D_1 sobre $\bar{D}$ que hace pasar puntos interiores a los interiores, puntos de frontera a los de frontera, con la particularidad de que dicha aplicación debe ser continuamente derivable y tener en D un jacobiano distinto de cero. Como siempre, dos representaciones se denominarán equivalentes, si se puede pasar de una de ellas a la otra con ayuda de una transformación admisible de los parámetros.

Diremos que la clase de representaciones equivalentes del tipo mencionado define una superficie continuamente derivable, siempre que en esta clase existe por lo menos una representación $r = r(u, v)$, $(u, v) \in \bar{D}$, que es continuamente derivable *hasta la frontera de la región* $\bar{D}$, es decir, definida en la clausura $\bar{D}$ de la región D.

Una superficie continuamente derivable se llama suave, si $r_u \times r_v \neq 0$ en $\bar{D}$ para cierta representación suya $r = r(u, v)$, $(u, v) \in \bar{D}$. El área de la superficie continuamente derivable representada por $r = r(u, v)$, $(u, v) \in \bar{D}$, se determina como valor de la integral

$$\iint_D |r_u \times r_v| \, du dv,$$

la cual es, quizás, impropia. Para convencerse de su existencia, basta realizar el cambio de variable con ayuda de una transformación admisible que convierte la representación dada en alguna otra que sea continuamente derivable hasta la frontera de la región.

De un modo semejante se debilitan los requisitos impuestos sobre las transformaciones admisibles del parámetro para el caso de las superficies orientadas.

Con estas definiciones quedan en vigor todas las definiciones aducidas anteriormente para las integrales de superficie, como también las propiedades de éstas, habida cuenta, naturalmente, de que en tal caso podemos obtener, para ciertas representaciones de las superficies, integrales impropias. Siguen siendo vigentes, además todos los teoremas concernientes a las integrales de superficie que se demuestran en el párrafo que sigue; no obstante, no nos detendremos en este caso especialmente.

Ejercicios. 2. Sea S una superficie suave en el nuevo sentido ampliado y sea Φ una función continua en S. Demuéstrese que existen las integrales

$$\int_S \Phi(x, y, z) dx dy, \quad \int_S \Phi(x, y, z) dz dx, \quad \int_S \Phi(x, y, z) dy dz.$$

Calcúlense las siguientes integrales de superficie de primera especie:

3. $\iint_S x^2 y^2 dS$; $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2, z \geqslant 0\}$.

4. $\iint_S \frac{dS}{x^2 + y^2 + z^2}$; $S = \{(x, y, z) : x^2 + y^2 = R^2, 0 \leqslant z \leqslant H\}$.

5. $\iint_S \frac{dS}{r}$; S es una porción de la superficie del paraboloide $z = xy$, obtenida al cortarla por el cilindro $x^2 + y^2 = R^2$, y r es la distancia del punto corriente de la superficie S hasta el eje Oz.

Calcúlense las siguientes integrales de superficie de segunda especie:

6. $\iint_S z dx dy$, donde S es el lado exterior del elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

7. $\iint_S yz dx dy + zx dy dz + xy dz dx$, donde S es el lado exterior de una superficie

compuesta por una porción de la superficie lateral del cilindro $x^2 + y^2 = R^2$ y las partes de los planos

$$x = 0, \quad y = 0, \quad z = 0, \quad z = H, \text{ siendo } \quad x, \quad y, \quad z \geqslant 0.$$

§ 52. CAMPOS ESCALARES Y VECTORIALES

52.1. DEFINICIONES

En lugar de los términos "función numérica de un punto", "función vectorial de un punto" se usan también términos equivalentes: *campo escalar, campo vectorial.* Esta terminología recalca que los valores de las funciones en consideración dependen precisamente de los puntos de un espacio (en los que dichas funciones están definidas) y no dependen de las coordenadas suyas, siendo elegido uno u otro sistema de coordenadas.

Empleando esta terminología, podemos decir, por ejemplo, que todo campo escalar $u = u(M)$, definido y derivable en cierta región G, engendra el campo vectorial de sus gradientes (véanse el p. 20.6 y el p. 50.5, pág 248): $a(M) = \operatorname{grad} u$.

Definición 1. *Supongamos que en la región* G [*]) *está dado un campo vectorial* $a = a(M)$ *y existe una función* $u = u(M)$, *definida en* G, *tal que* $a(M) = \operatorname{grad} u(M)$. *Entonces la función* $u(M)$ *lleva el nombre de función potencial o potencial del campo vectorial dado* [**]).

Introduciendo el símbolo nabla, $\nabla = i\dfrac{\partial}{\partial x} + j\dfrac{\partial}{\partial y} + k\dfrac{\partial}{\partial z}$ (véase el p. 20.7), podemos escribir:

$$\operatorname{grad} u = \nabla u,$$

donde en el segundo miembro figura el "producto" del vector simbólico nabla por una función numérica u.

Sea, por ejemplo, $E(M)$ la intensidad de un campo eléctrico creado por una carga unidad negativa que se ubica en el origen de coordenadas. Entonces en el punto $M(x, y, z)$ el vector $E(M)$ tiene, según lo sabemos por el curso de la física, la longitud $1/r^2$, donde $r = \sqrt{x^2 + y^2 + z^2}$, y está dirigido del punto M hacia el origen de coordenadas. De aquí se obtiene que

$$E(M) = \left(-\frac{x}{r^3}, -\frac{y}{r^3}, -\frac{z}{r^3}\right).$$

El potencial eléctrico del campo en consideración, es decir, la función $u(M) = 1/r$, es también potencial en el sentido indicado arriba, pues $\operatorname{grad} u(M) = E(M)$.

Consideraremos de nuevo un campo vectorial $a = a(M)$, definido en cierta región G. Fijemos un sistema de coordenadas y en este caso podemos considerar la función vectorial $a(M)$ como función de tres variables, esto es, las coordenadas x, y, z del punto M: $a = a(x, y, z)$.

---

[*]) En este párrafo, para simplificar, se considerarán sólo regiones planas o tridimensionales G.

[**]) A veces, en las aplicaciones, el potencial u se determina por la fórmula $a = -\operatorname{grad} u$.

Supongamos que $M_0 = (x_0, y_0, z_0) \in G$ y está dado un vector unidad $e = (\cos\alpha, \cos\beta, \cos\gamma)$. Tracemos por el punto M_0 una recta en la dirección de e:

$$x = x_0 + t\cos\alpha, \quad y = y_0 + t\cos\beta,$$
$$z = z_0 + t\cos\gamma, \quad -\infty < t < +\infty.$$

Definición 2. *La derivada de la función vectorial*

$$a(x_0 + t\cos\alpha, y_0 + t\cos\beta, z_0 + t\cos\gamma)$$

respecto de t para $t = 0$ *(si es que existe) recibe el nombre de derivada de la función vectorial* $a(M)$ *según la dirección de e en el punto* M_0 *y se denota* $\frac{\partial a}{\partial e}$:

$$\frac{\partial a(M_0)}{\partial e} = \frac{d}{dt}\, a(x_0 + t\cos\alpha, \quad y_0 = t\cos\beta, \quad z_0 + t\cos\gamma)\Big|_{t=0}.$$

Según la regla de derivación de una función compuesta, omitiendo, para simplificar, las designaciones del argumento, obtenemos

$$\frac{\partial a}{\partial e} = \frac{\partial a}{\partial x}\cos\alpha + \frac{\partial a}{\partial y}\cos\beta + \frac{\partial a}{\partial z}\cos\gamma. \tag{52.1}$$

Suponiendo $e\nabla = \cos\alpha\,\frac{\partial}{\partial x} + \cos\beta\,\frac{\partial}{\partial y} + \cos\gamma\,\frac{\partial}{\partial z}$ ("producto escalar del vector e por el vector simbólico ∇"), escribamos la fórmula (52.1) en la forma

$$\frac{\partial a}{\partial e} = (e\nabla)a.$$

Definición 3. *Si* $b = (b_x, b_y, b_z)$ *es un vector fijo arbitrario (no obligatoriamente un vector unidad), entonces el vector*

$$(b\nabla)a = b_x\frac{\partial a}{\partial x} + b_y\frac{\partial a}{\partial y} + b_z\frac{\partial a}{\partial z}$$

se llama gradiente del vector a respecto del vector b.

Si $b = bb_0$, donde $|b_0| = 1$, por medio de una "transformación formal" obtenemos

$$(b\nabla)a = (bb_0\nabla)a = b(b_0\nabla)a = b\,\frac{\partial a}{\partial b_0}.$$

Pasando a la notación en coordenadas, es fácil convencerse de que la fórmula obtenida es justa y mostrar que el símbolo ∇ se puede tratar en los cálculos como un vector auténtico sin olvidar, por supuesto, que ∇ significa, además, una operación de la derivación bien determinada. No nos detendremos aquí en la argumentación de legitimidad de tales "transformaciones formales con el símbolo ∇". Cualquier fórmula, obtenida de modo semejante, la podemos deducir, por supuesto, sin recurrir al símbolo ∇, sirviéndose de los razonamientos habituales argumentados en un sistema de coordenadas. Se debe tener en cuenta, sin embargo, que el empleo del símbolo ∇ en muchos casos facilita considerablemente los cálculos.

Volvamos otra vez al campo vectorial de partida $\boldsymbol{a} = (a_x, a_y, a_z)$ en la región G.

Definición 4. *Sea $\boldsymbol{a}(a_x, a_y, a_z)$ un campo derivable en cierto punto. El número $\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}$ se denomina divergencia del campo en el punto dado y se denota mediante div $\boldsymbol{a}$, es decir,*

$$\operatorname{div} \boldsymbol{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}. \tag{52.2}$$

Simbólicamente div $\boldsymbol{a}$ puede ser escrita como producto escalar del símbolo ∇ por el vector $\boldsymbol{a}$:

$$\operatorname{div} \boldsymbol{a} = \nabla \boldsymbol{a}.$$

Definición 5. *Un vector cuyas coordenadas son*

$$\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}, \quad \frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x}, \quad \frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \tag{52.3}$$

lleva el nombre de rotor del campo vectorial $\boldsymbol{a} = \boldsymbol{a}(M)$ y se designa rot $\boldsymbol{a}$.

Con ayuda del símbolo ∇ el rotor puede escribirse en forma del siguiente producto vectorial:

$$\operatorname{rot} \boldsymbol{a} = \nabla \times \boldsymbol{a} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix}. \tag{52.4}$$

El significado físico y geométrico de div $\boldsymbol{a}$ y rot $\boldsymbol{a}$ se aclarará en lo que sigue más abajo.

He aquí un ejemplo de las transformaciones formales con el símbolo ∇. Si el símbolo ∇ está seguido por varios términos y uno de los últimos se encuentra bajo la acción del símbolo en su calidad de operador de derivación, mientras que los otros términos están libres de la acción mencionada, el término accionado se denotará con una flecha vertical. Expliquemos esto con un ejemplo.

Sea f un campo escalar y $\boldsymbol{a}$, un campo vectorial, entonces

$$\operatorname{rot} f\boldsymbol{a} = \nabla \times f\boldsymbol{a} = \nabla \times f\overset{\downarrow}{\boldsymbol{a}} + \nabla \times f\overset{\downarrow}{\boldsymbol{a}} =$$

$$= f(\nabla \times \boldsymbol{a}) + (\nabla f \times \boldsymbol{a}) = f \operatorname{rot} \boldsymbol{a} + \operatorname{grad} f \times \boldsymbol{a}.$$

Introduzcamos unas definiciones más, ligadas con el campo vectorial $\boldsymbol{a} = (a_x, a_y, a_z)$ en la región G.

Definición 6. *Sea γ una curva cerrada suave a trozos en la región G. La integral*

$$\int_\gamma a_x dx + a_y dy + a_z dz$$

se denomina circulación del campo vectorial $\boldsymbol{a} = (a_x, a_y, a_z)$ a lo largo de la curva γ y se designa $\int_\gamma \boldsymbol{a} d\boldsymbol{r}$, donde $d\boldsymbol{r} = (dx, dy, dz)$.

Si γ es una curva suave orientada; $\boldsymbol{t} = (\cos\alpha, \cos\beta, \cos\gamma)$ es su vector tangente unidad; s, la longitud variable del arco y $\operatorname{pr}_{\boldsymbol{t}} \boldsymbol{a}$, la magnitud de la proyección del vec-

tor $\boldsymbol{a}$ sobre la tangente, entonces

$$\int_{\gamma} \boldsymbol{a}\,d\boldsymbol{r} = \int_{\gamma} \mathrm{pr}_{\boldsymbol{l}}\boldsymbol{a}\,ds.$$

Efectivamente,

$$\int_{\gamma} \boldsymbol{a}\,d\boldsymbol{r} = \int_{\gamma} a_x dx + a_y dy + a_z dz =$$

$$= \int_{\gamma} (a_x \cos\alpha + a_y \cos\beta + a_z \cos\gamma)ds = \int_{\gamma} at\,ds =$$

$$= \int_{\gamma} \mathrm{pr}_{\boldsymbol{l}}\boldsymbol{a}\,ds.$$

Definición 7. *Un campo cuya circulación a lo largo de cualquier curva cerrada suave a trozos, dispuesta en la región G, es nula, se denomina potencial.*

Recordemos que en el p. 47.8 se ha mostrado (véase el lema 2) que la condición, bajo la cual la integral $\int_{\gamma} P\,dx + Q\,dy$ es nula a lo largo de cualquier contorno cerrado $\gamma \subset G$, es equivalente a que $\int_{AB} P\,dx + Q\,dy$ no depende del trayecto de integración entre los puntos A y B. En la demostración de esta afirmación no se especificaba en ninguna circunstancia que la curva γ se dispone en una región plana. Por ello, la demostración del lema 2, aducida en el p. 47.8, queda vigente también para las integrales curvilíneas a lo largo de las curvas espaciales. De este modo, pues, *la circulación* $\int_{\gamma} \boldsymbol{a}\,d\boldsymbol{r} = \int_{\gamma} a_x dx + a_y dy + a_z dz$ *es nula a lo largo de cualquier contorno cerrado suave a trozos $\gamma \subset G$ cuando, y sólo cuando, la integral* $\int_{AB} a_x dx + a_y dy + a_z dz$ *no dependa del trayecto de integración*, es decir, de la curva con su origen en el punto A y extremo en el punto B y que está dispuesta íntegramente en la región G.

Consideremos a título de ejemplo un campo vectorial plano, es decir, el campo $\boldsymbol{a} = (P, Q)$ dado en la región plana G: $P = P(x, y)$, $Q = Q(x, y)$. El rotor de este campo tiene por expresión

$$\mathrm{rot}\,\boldsymbol{a} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & O \end{vmatrix} = \boldsymbol{k}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

En los términos nuevos el teorema 4 del p. 47.8 puede ser parafraseado de la manera siguiente. *Para una región plana simplemente conexa G son equivalentes la noción de carácter potencial del campo, la existencia de una función potencial y la condición de que el rotor del campo es nulo en todos los puntos.*

Definición 8. *Sea S una superficie orientada dispuesta en la región G, sea ν un vector unidad de la normal a la superficie que determina la orientación de la última,*

y sea S^+ la superficie S de orientación indicada. La integral

$$\iint\limits_{S^+} a_\nu dS$$

se llama flujo del campo vectorial a través de la superficie S y se designa

$$\iint\limits_{S} a dS,$$

donde

$$dS = \nu dS \text{ (o bien } \iint\limits_{S} a dS^+, dS^+ = \nu dS).$$

Es obvio que $a_\nu = \mathrm{pr}_\nu a$, razón por la cual $\iint\limits_{S} a dS = \iint\limits_{S} \mathrm{pr}_\nu a dS$.

En el flujo $\iint\limits_{S^+} a_\nu dS$ se omite corrientemente el índice de orientación y se escribe simplemente $\iint\limits_{S} a_\nu dS$, considerando que a título de orientación se ha tomado la normal ν que figura en la expresión subintegral.

En los puntos ulteriores de este párrafo se estudiarán algunas propiedades de los campos vectoriales, en particular, en el caso tridimensional se establecerán las condiciones necesarias y suficientes para que un campo sea potencial. Se demostrarán previamente los teoremas sobre las integrales, relacionados con los conceptos que se han introducido en el presente punto.

Ejercicios. 1. Demuéstrense las siguientes fórmulas:

a) rot grad $u = 0$;
b) div rot $a = 0$;
c) div grad $u = \Delta u$,
donde $\Delta u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}$;
d) rot rot a = grad div $a - \Delta a$, donde $\Delta a = (\Delta a_x, \Delta a_y, \Delta a_z)$, $a = (a_x, a_y, a_z)$;
e) div $(fa) = f$ div a + grad $f a$;
f) div $a \times b = b$ rot $a - a$ rot b.

2. Hállese el flujo del campo vectorial $a = (x - 2z)i + (x + 3y + z)j + (5x + y)k$ a través de una plazoleta triangular con vértices $A(1, 0, 0)$, $B(0, 1, 0)$, $C(0, 0, 1)$ y de orientación, determinada por una normal dirigida en el sentido opuesto al origen de coordenadas.

3. Hállese el flujo del campo vectorial $a = y^2 i + zk$ a través de la superficie S: $\{(x, y, z) : z = x^2 + y^2, 0 \leqslant z \leqslant 2\}$, si se conoce que la normal unitaria a la superficie citada está dirigida en el sentido opuesto al eje Oz.

4. Hállese el flujo del campo vectorial $a = xi + yj + \sqrt{x^2 + y^2 - 1}\,k$ a través de la porción de un hiperboloide $x^2 + y^2 - z^2 = 1$, $0 \leqslant z \leqslant \sqrt{3}$, opuesta al origen de coordenadas.

5. Hállese el flujo del campo vectorial $a = (xy - y^2)i + (-x^2 + xy + 2x)j + zk$ a través de un lado, opuesto al origen de coordenadas, de la parte de la superficie cilíndrica $x^2 + y^2 = 1$, que se obtiene al cortarla por el cono $z^2 = \dfrac{1}{2}x^2 + y^2$.

6. Hállese div a, si

$$a = (6x^2y^2 - z^3 + yz - 5)i + (4x^3y + xz + 2)j + (xy - 3xz^2 - 3)k.$$

Supongamos que $r = xi + yj + zk$, $r = |r|$, f es una función numérica derivable en todo punto de R_+, y c, un vector constante. Hállese:

7. div $(f(r)r)$.
8. div (rc)
9. div (grad $f(r)$).
10. div (r^2c).
11. div $(f(r)c)$.
12. div $(c \times r)$.
13. Hállese rot a, si

$$a = xyzi + (2x + 3y - z)j + (x^2 + z^2)k.$$

Hállese:

14. rot $(c \times r)$.
15. rot $((cr)r)$.
16. rot $(f(r)r)$.
17. rot $(f(r)c)$.

18. Hállese la circulación del campo vectorial $a = -yi + xj + ck$ (c = const) a lo largo de la circunferencia $(x - 2)^2 + y^2 = 1$, dispuesta en el plano Oxy.

19. Hállese la circulación del campo vectorial $a = yi - xj$ a lo largo de una línea cerrada que se forma por el arco de la astroide $x = R\cos^3 t$, $y = R\,\mathrm{sen}^3 t$, $(0 \leqslant t \leqslant \pi/2)$ y los segmentos de los ejes de coordenadas que se obtienen al cortarlos por la citada astroide.

52.2. SOBRE LA INVARIACIÓN DE LOS CONCEPTOS DE GRADIENTE, DIVERGENCIA Y ROTOR

Diremos en primer lugar que al transformar ortogonalmente las coordenadas cartesianas, el vector simbólico ∇ se transforma según las reglas de transformación de los vectores corrientes. En efecto, sea dada la transformación ortogonal de las coordenadas

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13}z, \\ y' &= a_{21}x + a_{22}y + a_{23}z, \\ z' &= a_{31}x + a_{32}y + a_{33}z. \end{aligned} \tag{52.5}$$

En el caso de las transfórmaciones ortogonales la matriz de una transformación inversa coincide con la matriz transpuesta, por lo cual

$$\begin{aligned} x &= a_{11}x' + a_{21}y' + a_{31}z' \\ y &= a_{12}x' + a_{22}y' + a_{32}z', \\ z &= a_{13}x' + a_{23}y' + a_{33}z'. \end{aligned} \tag{52.6}$$

Con ello, como es bien conocido, a partir de las fórmulas (52.5) y (52.6) se transforman tanto las coordenadas de los puntos, como las de los vectores.

Haciendo uso de las fórmulas (52.5) y la regla de derivación de una función compuesta, obtendremos

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial}{\partial y'}\frac{\partial y'}{\partial x} + \frac{\partial}{\partial z'}\frac{\partial z'}{\partial x} =$$

$$= a_{11}\frac{\partial}{\partial x'} + a_{21}\frac{\partial}{\partial y'} + a_{31}\frac{\partial}{\partial z'},$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial x'}\frac{\partial x'}{\partial y} + \frac{\partial}{\partial y'}\frac{\partial y'}{\partial y} + \frac{\partial}{\partial z'}\frac{\partial z'}{\partial y} =$$
$$= a_{12}\frac{\partial}{\partial x'} + a_{22}\frac{\partial}{\partial y'} + a_{32}\frac{\partial}{\partial z'}, \qquad (52.7)$$
$$\frac{\partial}{\partial z} = \frac{\partial}{\partial x'}\frac{\partial x'}{\partial z} + \frac{\partial}{\partial y'}\frac{\partial y'}{\partial z} + \frac{\partial}{\partial z'}\frac{\partial z'}{\partial z} =$$
$$= a_{13}\frac{\partial}{\partial x'} + a_{23}\frac{\partial}{\partial y'} + a_{33}\frac{\partial}{\partial z'}$$

Correspondientemente, las fórmulas inversas que expresan derivadas respecto de las variables x', y', z' en términos de las derivadas respecto de x, y, z, tendrán por expresión

$$\frac{\partial}{\partial x'} = a_{11}\frac{\partial}{\partial x} + a_{12}\frac{\partial}{\partial y} + a_{13}\frac{\partial}{\partial z},$$
$$\frac{\partial}{\partial y'} = a_{21}\frac{\partial}{\partial x} + a_{22}\frac{\partial}{\partial y} + a_{23}\frac{\partial}{\partial z}, \qquad (52.8)$$
$$\frac{\partial}{\partial z'} = a_{31}\frac{\partial}{\partial x} + a_{32}\frac{\partial}{\partial y} + a_{33}\frac{\partial}{\partial z}.$$

Las fórmulas (52.5) — (52.8) muestran precisamente que las coordenadas de los vectores ordinarios y las "coordenadas" del vector simbólico ∇ se transforman, en las transformaciones ortogonales de coordenadas cartesianas, según una misma regla. En particular, de (52.8) se deduce que el gradiente de la función u en el sistema de coordenadas x, y, z, es decir, el vector de coordenadas $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}$ tendrá en el sistema x', y', z' las coordenadas $\frac{\partial u}{\partial x'}, \frac{\partial u}{\partial y'}, \frac{\partial u}{\partial z'}$, es decir, sigue siendo gradiente también en este último sistema de coordenadas. De este modo, queda demostrado una vez más (véase el p. 20.7) que el gradiente de una función no depende de la elección del sistema de coordenadas cartesianas. Por cuanto el vector ∇ se transforma igual que los vectores ordinarios, es natural esperar que el producto escalar ∇a tampoco depende de la elección del sistema mencionado de coordenadas.

Supongamos que el vector a en el sistema x, y, z tiene las coordenadas a_x, a_y, a_z y en el sistema x', y', z', las $a_{x'}$, $a_{y'}$, $a_{z'}$. En virtud de las fórmulas (52.7), tenemos

$$\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} = a_{11}\frac{\partial a_x}{\partial x'} + a_{21}\frac{\partial a_x}{\partial y'} + a_{31}\frac{\partial a_x}{\partial z'} +$$
$$+ a_{12}\frac{\partial a_y}{\partial x'} + a_{22}\frac{\partial a_y}{\partial y'} + a_{32}\frac{\partial a_y}{\partial z'} + a_{13}\frac{\partial a_z}{\partial x'} + a_{23}\frac{\partial a_z}{\partial y'} +$$
$$+ a_{33}\frac{\partial a_z}{\partial z'} = \frac{\partial}{\partial x'}(a_{11}a_x + a_{12}a_y + a_{13}a_z) +$$

$$+\frac{\partial}{\partial y'}(a_{21}a_x + a_{22}a_y + a_{23}a_z) + \frac{\partial}{\partial z'}(a_{31}a_x + a_{32}a_y + a_{33}a_z). \quad (52.9)$$

Aplicando las fórmulas (52.5) al vector $a = (a_x, a_y, a_z)$, (es decir, sustituyendo en dichas fórmulas x, y, z por a_x, a_y, a_z, y x', y', z', por $a_{x'}, a_{y'}, a_{z'}$), llegamos a que las expresiones entre paréntesis en el segundo miembro de la igualdad (52.9) son iguales sucesivamente a $a_{x'}$, $a_{y'}$, $a_{z'}$, y, por consiguiente,

$$\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} = \frac{\partial a_{x'}}{\partial x'} + \frac{\partial a_{y'}}{\partial y'} + \frac{\partial a_{z'}}{\partial z'}.$$

Esta igualdad muestra precisamente que la divergencia de un campo vectorial en todo punto se determina unívocamente por el propio campo vectorial y no depende de cómo se elige el sistema de coordenadas, lo que al principio podría parecer proveniente de la fórmula (52.2).

Un producto vectorial de los vectores ordinarios no depende, en virtud de su significado geométrico, de la elección de los sistemas cartesianos de coordenadas con la misma orientación (por ejemplo, el producto vectorial de dos vectores no cambiará, si se pasa de un sistema derecho de coordenadas cartesianas (véase el p. 50.8) a otro sistema de la misma orientación. Es por eso que resulta natural esperar que la misma propiedad la posee también el "producto vectorial simbólico" rot $a = \nabla \times a$.

Efectivamente, si designamos los vectores coordenados unidad del sistema de coordenadas x', y', z' mediante i', j', k', respectivamente, entonces, como se sabe, los vectores coordenados unidad i, j, k del sistema de coordenadas x, y, z se expresarán en términos de i', y', k' mediante una matriz transpuesta respecto de la matriz de la transformación (52.5), es decir, mediante la matriz de la transformación (52.6):

$$\begin{aligned} i &= a_{11}i' + a_{21}j' + a_{31}k', \\ j &= a_{12}i' + a_{22}j' + a_{32}k', \\ k &= a_{13}i' + a_{23}j' + a_{33}k'. \end{aligned} \quad (52.10)$$

Haciendo uso de las fórmulas (52.6), (52.7) y (52.10), obtenemos

$$\text{rot } a = \nabla \times a = \begin{vmatrix} i & j & k \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11}i' + a_{21}j' + a_{31}k' & a_{12}i' + a_{22}j' + a_{32}k' & a_{13}i' + a_{23}j' + a_{33}k' \\ a_{11}\dfrac{\partial}{\partial x'} + a_{21}\dfrac{\partial}{\partial y'} + a_{31}\dfrac{\partial}{\partial z'} & a_{12}\dfrac{\partial}{\partial x'} + a_{22}\dfrac{\partial}{\partial y'} + a_{32}\dfrac{\partial}{\partial z'} & a_{13}\dfrac{\partial}{\partial x'} + a_{23}\dfrac{\partial}{\partial y'} + a_{33}\dfrac{\partial}{\partial z'} \\ a_{11}a_{x'} + a_{21}a_{y'} + a_{31}a_{z'} & a_{12}a_{x'} + a_{22}a_{y'} + a_{32}a_{z'} & a_{13}a_{x'} + a_{23}a_{y'} + a_{33}a_{z} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \cdot \begin{vmatrix} i' & j' & k' \\ \dfrac{\partial}{\partial x'} & \dfrac{\partial}{\partial y'} & \dfrac{\partial}{\partial z'} \\ a_{x'} & a_{y'} & a_{z'} \end{vmatrix}. \qquad (52.11)$$

La última igualdad se demuestra igual que en el caso de las matrices numéricas ordinarias se demuestra el hecho de que el determinante de un producto de dos matrices cuadradas de un mismo orden es igual al producto de sus determinantes. Con el fin de demostrar dicha igualdad, basta convencerse de que en sus ambos miembros figuran las sumas algebraicas iguales de los mismos sumandos.

El determinante de la transformación ortogonal es igual a $+1$ ó -1, con la particularidad de que si dicha transformación conserva su orientación, entonces a $+1$. Por ello, si en el caso que se considera los sistemas elegidos de coordenadas x, y, z y x', y', z' están igualmente orientados, tendremos

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 1$$

y, por consiguiente, de (52.11) se tiene

$$\begin{vmatrix} i & j & k \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix} = \begin{vmatrix} i' & j' & k' \\ \dfrac{\partial}{\partial x'} & \dfrac{\partial}{\partial y'} & \dfrac{\partial}{\partial z'} \\ a_{x'} & a_{y'} & a_{z'} \end{vmatrix}$$

Esta igualdad significa precisamente que el rotor del campo vectorial no depende de cómo se elige el sistema de coordenadas cartesianas con la misma orientación que tiene el sistema dado. Indiquemos, sin embargo, si se pasa de un sistema de coordenadas al otro sistema de otra orientación, por ejemplo, de un sistema de coordenadas defecho a un sistema izquierdo, entonces todo rotor (al igual que un producto vectorial ordinario) se sustituirá por un vector opuesto. Esto se deduce de la fórmula (52.11), puesto que el determinante de la transformación ortogonal, que cambia la orientación, es igual a -1.

De este modo, el rotor de un campo vectorial se determina univocamente, "salvo el signo", por el propio campo vectorial, y si no limitamos únicamente a los sistemas derechos de coordenadas cartesianas, no depende de la elección de los mismos.

52.3. FÓRMULA DE OSTROGRADSKI — GAUSS. DEFINICIÓN GEOMÉTRICA DE LA DIVERGENCIA

Sea G una región en el espacio R^3_{xyz}. Supongamos que en el plano R^2_{xy} existe tal dominio cuadrable Γ que la frontera de la región G se compone de dos superficies S_1 y S_2, definidas mediante las representaciones explícitas $z = \varphi(x, y)$ y $z = \psi(x, y)$, respectivamente (donde las funciones $\varphi(x, y)$ y $\psi(x, y)$ son continuas en la región cerrada $\overline{\Gamma}$, $\varphi(x, y) \leqslant \psi(x, y)$, $(x, y) \in \overline{\Gamma}$), y, quizás, de una parte S_0 del cilindro cuya base está formada por la frontera $\partial\Gamma$ de la región Γ (véase el p. 44.1). Supongamos, además, que S_0, S_1 y S_2 son superficies suaves a trozos (fig. 222). En este caso toda la frontera S de la región G será asimismo una superficie suave a trozos y, además, orientable, como cualquier superficie suave a trozos que sirve de frontera de una región. Las normales exteriores ν de la superficie S en sus partes suaves constituyen las orientaciones de las mismas. Debido a estas orientaciones, las partes suaves de la superficie S están orientadas de manera concordante (véase el p. 50.11) y, por consiguiente, generan la orientación de toda la superficie S. Esta orientación se logra, si para cada parte suave de la superficie se elige la orientación de su borde, concordada con la normal exterior ν en la parte respectiva según la regla del sacacorchos.

Designemos la superficie S y, correspondientemente, las superficies S_0, S_1 y S_2 de orientación elegida (la que se llamará positiva) mediante S^+ y, correspondientemente, S_0^+, S_1^+, S_2^+. Observemos que aquí de orientación positiva sirve para la superficie S_1 su "cara inferior", y para la superficie S_2, la "cara superior" de ella (véase 51.1).

La elección de la normal ν en el caso que se considera se describe también de una manera fácil y directa, es decir, sin que se recurra al concepto de normal "exterior": en los puntos de la superficie S_1, en los que la normal existe, se debe elegir una normal que forma el ángulo obtuso con el eje Oz, y en los puntos de la superficie S_2, un ángulo agudo. Entre tanto, en los puntos de la superficie S_0 la elección de la normal

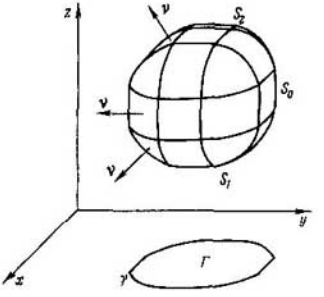


Fig. 222

para nuestros objetivos no nos importa (lo que se pondrá claro en adelante): calcularemos las integrales del tipo (51.7) extendidas a la superficie S, las cuales, siendo extendidas a la superficie S_0, son nulas, cualquiera que sea la elección de las normales, puesto que las últimas son siempre perpendiculares al eje Oz.

Supondremos que la región G posee las propiedades, análogas a las citadas, respecto de todos los ejes coordenados. Tales *regiones* se llamarán *elementales* (compárese con el p. 47.5). Un ejemplo de región elemental se expone en la fig. 222.

Designemos con $\cos\alpha$, $\cos\beta$, $\cos\gamma$ los cosenos directores de la normal exterior unitaria ν de la superficie S:

$$\nu = (\cos\alpha, \cos\beta, \cos\gamma).$$

Teorema 1 (de Ostrogradski — Gauss *). *Supongamos que en la clausura $\overline{G}$ de la región G, del tipo mencionado más arriba, vienen dadas las funciones $P = P(x, y, z)$, $Q = Q(x, y, z)$ y $R = R(x, y, z)$ continuas en $\overline{G}$, lo mismo que todas las derivadas parciales suyas $\frac{\partial P}{\partial x}, \frac{\partial Q}{\partial y}, \frac{\partial R}{\partial x}$.* **) *En este caso*

$$\iiint_G \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dxdydz =$$

$$= \iint_{S_+} (P\cos\alpha + Q\cos\beta + R\cos\gamma)dS. \qquad (52.12)$$

Esta fórmula, al suponer $a = (P, Q, R)$, puede ser escrita en la forma

$$\iiint_G \operatorname{div} a\, dxdydz = \iint_S a\, dS^+, \qquad (52.13)$$

es decir, la integral de la divergencia de un campo vectorial, extendida a cierta región, es igual al flujo de este campo a través de la superficie que limita dicha región.

DEMOSTRACIÓN. Consideremos, por ejemplo, la integral

$$\iiint_G \frac{\partial R}{\partial z}\, dxdydz.$$

Al emplear las designaciones introducidas al principio de este punto, obtendremos

$$\iiint_G \frac{\partial R}{\partial z}\, dxdydz = \iint_{\Gamma} \left[\int_{\varphi(x,y)}^{\psi(x,y)} \frac{\partial R}{\partial z}\, dz\right] dxdy =$$

*) M. V. Ostrogradski (1801—1861), matemático ruso; K. F. Gauss (1777—1855), matemático alemán.

**) La continuidad de las derivadas parciales en una frontera se entiende como su prolongabilidad continua a la frontera de la región.

$$= \iint_{\Gamma} [R[x, y, \psi(x, y)] - R[x, y, \varphi(x, y)]]dxdy =$$

$$= \iint_{S_2^+} R(x, y, z)dxdy + \iint_{S_1^+} R(x, y, z)dxdy. \quad (52.14)$$

Al observar luego que en la superficie S_0 tiene lugar la igualdad $\cos\gamma = 0$, tendremos (véase (51.7))

$$\iint_{S_0^+} R(x, y, z)dxdy = \iint_{S_0^+} P(x, y, z)\cos\gamma dS = 0.$$

Por eso la fórmula (52.14) puede ser escrita así:

$$\iiint_{G} \frac{\partial R}{\partial z} dxdydz =$$

$$= \iint_{S_2^+} Rdxdy + \iint_{S_1^+} Rdxdy + \iint_{S_0^+} Rdxdy = \iint_{S^+} Rdxdy. \quad (52.15)$$

De un modo sumamente análogo se demuestran las fórmulas

$$\iiint_{G} \frac{\partial P}{\partial x} dxdydz = \iint_{S^+} Pdydz,$$

$$\iiint_{G} \frac{\partial Q}{\partial y} dxdydz = \iint_{S^+} Qdzdx. \quad (52.16)$$

Sumando (52.15) y (52.16), obtendremos precisamente, en virtud de las definiciones (51.7) y (51.12), la fórmula (52.12), llamada *fórmula de Ostrogradski — Gauss*. □

A veces resulta más conveniente utilizar la fórmula (52.12) en la forma

$$\iiint_{G} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial x}\right) dxdydz = \iint_{S} Pdydz + Qdzdx + Rdxdy.$$

La validez de tal notación se deduce directamente de la definición de la integral de superficie de segunda especie: véase (51.7) y (51.12).

La fórmula de Ostrogradski — Gauss (52.12) puede demostrarse también para las regiones G de la forma más general en comparación con las mencionadas, a saber, para aquellas que admiten la partición finita en las regiones G_i, $i = 1, 2, \ldots, i_0$, del tipo considerado más arriba. Con este fin es suficiente escribir la fórmula de Ostrogradski para cualquiera de las regiones G_i y sumar los resultados; de resultas se obtiene la fórmula buscada para la región G. Efectivamente, en el primer miembro de la igualdad se obtendrá, en virtud de que la integral es aditiva, una in-

tegral correspondiente extendida a la región G, mientras que en el segundo miembro las integrales de superficie calculadas a lo largo de las partes correspondientes de las fronteras de las regiones G_i darán en suma cero, porque las normales exteriores en los puntos de fronteras de las regiones G_i, pertenecientes a las fronteras de dos regiones de este género, están dirigidas en las direcciones opuestas; quedarán, pues, sólo las integrales a lo largo de aquellas partes de las fronteras de G_i que forman en su totalidad la frontera de la región G (compárese con el p. 47.5). Los planos paralelos a los planos coordenados suelen ser más cómodos para obtener las particiones citadas de la región G.

Observemos que entre las regiones de este género hay también aquellas cuya frontera se compone de unos cuantos "trozos", es decir, puede ser representada como suma de un número finito de superficies disjuntas suaves a trozos (compárese con las generalizaciones correspondientes de la fórmula de Green en el p. 47.5).

Se puede mostrar que la fórmula de Ostrogradski — Gauss es válida para cualquier región limitada cuya frontera se compone de un número finito de las superficies suaves a trozos. Sin embargo, esto sería un trabajo muy voluminoso y no nos detendremos en dicho problema, limitándonos solamente a la formulación del teorema.

Teorema 1 (de Ostrogradski — Gauss). *Supongamos que la frontera ∂G de una región limitada G se compone de un número finito de superficies suaves a trozos y que el vector $a = (P, Q, R)$, como también las derivadas parciales $\frac{\partial P}{\partial x}, \frac{\partial Q}{\partial y}$ y $\frac{\partial R}{\partial z}$ son continuas en $\bar{G}$, entonces*

$$\iiint_G \operatorname{div} a\,dx\,dy\,dz = \iint_{\partial G} a\,dS.$$

Como orientación en las partes suaves de la frontera ∂G está elegida aquí la normal exterior.

Por ejemplo, si $G = \{(x, y, z) : 0 < a < \sqrt{x^2 + y^2 + z^2} < b\}$ es un anillo esférico y, por lo tanto, su frontera se compone de dos esferas

$$S_1 = \{(x, y, z) : x^2 + y^2 + z^2 = a^2\} \quad \text{y}$$
$$S_2 = \{(x, y, z) : x^2 + y^2 + z^2 = b^2\},$$

entonces, en la esfera interior S_1 se debe tomar una normal dirigida hacia el centro de la bola G y en la esfera exterior S_2, en la dirección contraria.

La fórmula de Ostrogradski — Gauss permite hallar la expresión para el volumen de una región mediante la integral correspondiente de superficie. Efectivamente, poniendo en (52.12), $P(x, y, z) = x$, $Q(x, y, z) = y$, $R(x, y, z) = z$, y observando que $\iiint_G dx\,dy\,dz = \mu G$, obtendremos

$$\mu G = \frac{1}{3} \iint_{S+} (x \cos\alpha + y \cos\beta + z \cos\gamma)\,dS,$$

o bien

$$\mu G = \frac{1}{3} \iint_{S+} x\,dy\,dz + y\,dz\,dx + z\,dx\,dy.$$

La fórmula de Ostrogradski — Gauss ofrece la posibilidad de establecer un enfoque geométrico al concepto de divergencia.

Teorema 2. *Supongamos que en una región tridimensional G *) está definido el campo vectorial continuamente derivable $\boldsymbol{a} = \boldsymbol{a}(M)$. Sea $M_0 \in G$ y sea D tal región con la frontera suave a trozos S que $M_0 \in D$, $\bar{D} \subset G$ y para el dominio D es válida la fórmula de Ostrogradski — Gauss* **).

Designemos con S^+ la superficie S, orientada con ayuda de la elección de la normal exterior, y con $d(D)$, el diámetro del dominio D. Entonces

$$\operatorname{div} \boldsymbol{a}(M_0) = \lim_{d(D)\to 0} \frac{\iint_S \boldsymbol{a}\, dS^+}{\mu D}. \tag{52.17}$$

DEMOSTRACIÓN. Según la fórmula (52.13) tenemos

$$\iiint_D \operatorname{div} \boldsymbol{a}\, dx\, dy\, dz = \iint_S \boldsymbol{a}\, dS^+. \tag{52.18}$$

Mas según el teorema integral del valor medio (p. 44.6),

$$\iiint_D \operatorname{div} \boldsymbol{a}\, dx\, dy\, dz = \operatorname{div} \boldsymbol{a}(M)\mu D, \quad M \in D. \tag{52.19}$$

Al sustituir (59.19) en (52.18), obtenemos

$$\operatorname{div} \boldsymbol{a}(M) = \frac{\iint_S \boldsymbol{a}\, dS^+}{\mu D}. \tag{52.20}$$

Pasando al límite en la fórmula (52.20) para $d(D) \to 0$, obtendremos, por ser continua la función $\operatorname{div} \boldsymbol{a}(M)$ en el punto M_0, la fórmula (52.17). □.

Se puede probar que las magnitudes en el segundo miembro de la igualdad (52.27) no dependen de la elección del sistema de coordenadas (en el segundo miembro figuran la integral doble de un producto escalar de los vectores y el volumen de la región), por lo cual de aquí se deduce una vez más que la divergencia del campo vectorial no depende de la elección del sistema de coordenadas.

De la igualdad (52.17) se infiere que el segundo miembro de esta igualdad puede considerarse como la definición de la divergencia del campo dado.

Los puntos del campo vectorial $\boldsymbol{a}$, en los que $\operatorname{div} \boldsymbol{a} \neq 0$, reciben el nombre de "fuentes" del campo vectorial. El carácter natural de este término se explica intuitivamente por aquella circunstancia que si el punto es una "fuente", entonces, como se deduce de la fórmula (52.17), para todas las regiones D, cuyos diámetros son suficientemente pequeños y que contienen el punto M_0, tendremos $\iint_S \boldsymbol{a}\, dS \neq 0$, es decir, el flujo a través de cualquier superficie, suficientemente pequeña, que rodea la fuente, no es igual a cero.

---

*) Aquí no se imponen ningunas restricciones sobre la estructura de la región G.

**) Las regiones D de esta índole siempre existen, por ejemplo, entre ellas figuran todas las bolas de radio suficientemente pequeño con centro en el punto M_0, o bien los cubos de dimensiones suficientemente pequeñas con centro en el punto M_0.

52.4. FÓRMULA DE STOKES. DEFINICIÓN GEOMÉTRICA DEL ROTOR

Supongamos que S es una superficie en el espacio R^3_{xyz} dos veces continuamente derivable privada de puntos singulares; $r = r(u, v)$, $(u, v) \in \bar{D}$ es su representación y D, una región limitada plana, para la cual es válida la fórmula de Green. Admitamos que la frontera de la región D se compone de un contorno sencillo suave a trozos. Designemos con γ_0 el contorno orientado positivo que limita el dominio D, y con $u = u(t)$, $v = v(t)$, $a \leqslant t \leqslant b$, la representación de este contorno. Sea

$$\nu = \frac{r_u \times r_v}{|r_u \times r_v|}$$

la orientación en la superficie S (véase la definición 20 en el p. 50.8), $\nu = (\cos\alpha, \cos\beta, \cos\gamma)$. Bajo las suposiciones asumidas la normal ν es continua en $\bar{D}$.

Designemos mediante S^+ la superficie S con la normal ν elegida en S. Sea Γ un contorno que tiene la representación $r = r(u(t), v(t))$, $a \leqslant t \leqslant b$. Diremos que el *contorno* Γ *limita la superficie* S *y también que la superficie* S *está tendida sobre el contorno* Γ.

Supongamos, por fin, que G es una región en el espacio R^3_{xyz} y $S \subset G$. Cumplidas estas suposiciones, resulta válido el siguiente teorema.

Teorema 3 (de Stokes *)). *Supongamos que las funciones* P, Q *y* R *son continuas en la región* G, *al igual que sus primeras derivadas parciales, y sea* $a = (P, Q, R)$. *En este caso*

$$\int_\Gamma a\,dr = \iint_S \operatorname{rot} a\,dS^+, \tag{52.21}$$

es decir, la circulación de un campo vectorial a lo largo del contorno Γ *es igual al flujo del rotor de este campo a través de la superficie* S, *limitada por el contorno* Γ. *En la forma coordenada esta fórmula tiene por expresión*

$$\int_\Gamma P\,dx + Q\,dy + R\,dz = \iint_S \begin{vmatrix} \cos\alpha & \cos\beta & \cos\gamma \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} dS,$$

o bien

$$\int_\Gamma P\,dx + Q\,dy + R\,dz = \iint_S \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\cos\beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\cos\gamma\right] dS. \tag{52.22}$$

DEMOSTRACIÓN. Examinemos, por ejemplo, la integral $\int_\Gamma P\,dx$. Al observar que a lo largo de las curvas Γ_0 y Γ las variables u y v son funciones de t, y al hacer uso de las designaciones introducidas al principio de este punto, obtendremos

*) G. Stokes (1819—1903), matemático y mecánico inglés.

$$\int_{\Gamma} P(x, y, z)dx = \int_{a}^{b} P[x(u(t), v(t)), y(u(t), v(t)),$$

$$z(u(t), v(t))]x'_t(u(t), v(t))dt = \int_{\Gamma_0} P[x(u, v), y(u, v),$$

$$z(u, v)]\left[\frac{\partial x(u, v)}{\partial u}du + \frac{\partial x(u, u)}{\partial v}dv\right].$$

Hemos aprovechado aquí la fórmula

$$x'_t(u(t), v(t)) = \frac{\partial x(u(t), v(t))}{\partial u}\frac{\partial u}{dt} + \frac{\partial x(u(t), v(t))}{\partial v}\frac{\partial v}{dt}.$$

Al aplicar la fórmula de Green a la integral obtenida $\int_{\Gamma_0} P\frac{\partial x}{\partial u}du + P\frac{\partial x}{\partial v}dv$, tendremos:

$$\int_{\Gamma} Pdx = \iint_{D}\left[\frac{\partial}{\partial u}\left(P\frac{\partial x}{\partial v}\right) - \frac{\partial}{\partial v}\left(P\frac{\partial x}{\partial u}\right)\right]dudv =$$

$$= \iint_{D}\left[\left(\frac{\partial P}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial P}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial u}\right)\frac{\partial x}{\partial v} + P\frac{\partial^2 x}{\partial u \partial v} -\right.$$

$$\left. - \left(\frac{\partial P}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial P}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial v}\right)\frac{\partial x}{\partial u} - P\frac{\partial^2 x}{\partial v \partial u}\right]dudv =$$

$$= \iint_{D}\left[\frac{\partial P}{\partial z}\frac{\partial(z, x)}{\partial(u, v)} - \frac{\partial P}{\partial y}\frac{\partial(x, y)}{\partial(u, v)}\right]dudv =$$

$$= \iint_{S^+}\frac{\partial P}{\partial z}dzdx - \iint_{S^+}\frac{\partial P}{\partial y}dxdy = \iint_{S}\left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right)dS. \quad (52.23)$$

Aquí se han tomado en consideración las fórmulas (51.7), (51.10) y (51.13). Análogamente se demuestra que

$$\int_{\Gamma} Qdy = \iint_{S}\left(\frac{\partial Q}{\partial x}\cos\gamma - \frac{\partial Q}{\partial z}\cos\alpha\right)dS, \quad (52.24)$$

$$\iint_{\Gamma} Rdz = \iint_{S}\left(\frac{\partial R}{\partial y}\cos\alpha - \frac{\partial R}{\partial x}\cos\beta\right)dS. \quad (52.25)$$

Sumando las fórmulas (52.23), (52.24) y (52.25), obtendremos, precisamente la *fórmula* (52.22) que lleva el nombre *de Stokes.* □

Para imaginarse de la manera más ilustrativa la relación que existe entre la elección de la normal ν en la superficie S y la orientación del contorno Γ, *que la limita*, consideraremos una superficie S que tiene la representación explícita $z = f(x, y)$, $(x, y) \in \bar{D}$.

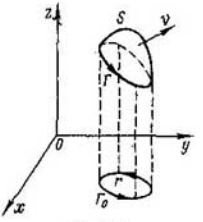


Fig. 223

Sea Γ_0 un contorno orientado positivo en el plano xOy que es al mismo tiempo la frontera de D, y sea $x = x(t)$, $y = y(t)$, $a \leqslant t \leqslant b$, la representación de dicho contorno. La orientación de la curva Γ la definamos, como hasta ahora, mediante la representación

$$x = x(t), \quad y = y(t), \quad z = f[x(t), y(t)], \quad a \leqslant t \leqslant b. \tag{52.26}$$

En el caso dado el contorno Γ_0 es la proyección de la curva Γ. Entre tanto, la normal ν, como ya se ha mostrado, forma con el eje Oz (cuando la superficie se representa de manera explícita) un ángulo agudo (véase el p. 51.1), por lo cual si la superficie S se observa desde la dirección positiva del eje Oz, el contorno Γ quedará orientado en el sentido contrahorario, es decir, la orientación de la curva Γ está concordada con la normal ν ("según la regla del sacacorchos") (fig. 223). Esto es equivalente a que un observador, que recorre la superficie S a lo largo del contorno orientado Γ y que la mira desde el extremo de la normal ν, ve la superficie S dispuesta a la izquierda. Esta interpretación evidente de concordancia entre la orientación de la normal ν y la del contorno Γ tiene la ventaja de que no está relacionada con la elección del sistema de coordenadas y queda válida para cualquier superficie S, considerada en relación con el teorema de Stokes, y no sólo para una superficie dada explícitamente. Por supuesto, todos semejantes razonamientos no representan demostraciones matemáticas, sino que sólo explican de modo palpable la fórmula de Stokes.

Cabe observar que la fórmula de Stokes queda válida, si consideramos en ella la orientación opuesta del contorno y las normales opuestas $-\nu$; en este caso ambos miembros de la igualdad (52.21) cambian el signo por el contrario (las orientaciones del contorno y de la superficie quedan concordadas según la "regla del sacacorchos").

La fórmula de Stokes puede ser demostrada también para las superficies orientables suaves a trozos $S = \{S_i\}_{i=1}^{i_0}$, a saber, en el caso cuando las superficies S_i, $i = 1, 2, \ldots, i_0$, satisfacen las condiciones del teorema demostrado 3. En este último caso el borde de la superficie ∂S (véase el p. 50.11) puede ser de un número finito de los contornos cerrados Γ_j, $j = 1, 2, \ldots, j_0$.

Para demostrarlo, basta escribir las fórmulas de Stokes para cada superficie S_i, $i = 1, 2, \ldots, i_0$, y sumarlas (compárese con las generalizaciones de la fórmula de

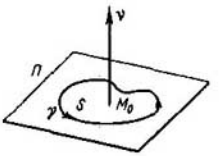


Fig. 224

Green en el p. 47.5 y del teorema de Ostrogradski — Gauss en el n. 52.3).

Indiquemos, además, que la condición de que la superficie S sea *dos veces* continuamente derivable fue impuesta en el teorema 3 sólo con el objeto de simplificar la demostración (la última se hace *esencialmente* más simple). La fórmula de Stokes (52.21) queda en vigor también, si se supone sólo la suavidad de la superficie S (conservándose las demás condiciones del teorema). La demostración de esta afirmación sale de los márgenes de nuestro curso.

De lo dicho proviene que la fórmula de Stokes sigue siendo lícita también para las superficies $S = \{S_i\}_{i=1}^{i=i_0}$ que son simplemente orientadas y suaves a trozos (es decir, sin la suposición de que las superficies S_i hayan de ser *dos veces* continuamente derivables).

Enunciemos ahora el teorema para este caso.

Teorema 3′ (de Stokes). *Supongamos que la función vectorial a es continuamente derivable en la región G y sea $S = \{S_i\}_{i=1}^{i=i_0}$ una superficie orientada suave a trozos cuyo borde ∂S tiene orientación engendrada por la orientación dada de la superficie S* (véase el p. 50.11). *Entonces*

$$\int_{\partial S} a\,dr = \iint_S \operatorname{rot} a\,dS.$$

La concordancia de las orientaciones de los contornos Γ_j, de los que se compone el borde ∂S de la superficie S, con la orientación de dicha superficie y, por lo tanto, con la de ν de las superficies S_i significa con evidencia que un observador, que se mueve a lo largo del contorno Γ_j, $j = 1, 2 \dots, j_0$, y mira a la superficie S desde la punta de la normal ν, ve la superficie S dispuesta a la izquierda.

El teorema de Stokes ofrece la posibilidad de establecer un enfoque geométrico al concepto de rotor de un campo vectorial.

Teorema 4. *Supongamos que en la región tridimensional G está definido un campo vectorial continuamente derivable $a = a(M)$; M_0 es un punto fijo, $M_0 \in G$, ν es un vector unidad constante arbitrariamente elegido y Π, un plano que es perpendicular a ν y pasa por el punto M_0; S es una región limitada en el plano Π cuya frontera está representada por el contorno suave a trozos Γ, $d(S)$ es el diámetro de la región S; supongamos también que el contorno Γ está orientado, de modo con-*

cordado, con la normal ν **), $M_0 \in S$ y $S \subset G$ **) (fig. 224). En este caso **\)*

$$\operatorname{rot}_\nu a(M_0) = \lim_{d(S)\to 0}\left(\frac{\int\limits_\Gamma a\,dr}{\mu S}\right). \tag{52.27}$$

DEMOSTRACIÓN. Según la fórmula de Stokes,

$$\int\limits_\Gamma a\,dr = \iint\limits_S rot_\nu a\,dS,$$

mas, conforme al teorema integral del valor medio,

$$\iint\limits_S \operatorname{rot}_\nu a\,dS = \operatorname{rot}_\nu a(M)\mu S, \quad M \in S.$$

Por consiguiente,

$$\operatorname{rot}_\nu a(M) = \frac{\int\limits_\Gamma a\,dS}{\mu S}. \tag{52.28}$$

Observemos que cuando $d(S) \to 0$, también $M \to M_0$. En vista de que la función $\operatorname{rot}_\nu a(M)$ es continua en el punto M_0, pasando al límite en (52.28) para $d(S) \to 0$, obtendremos la fórmula (52.27). □

De (52.27) se deduce que el segundo miembro de esta igualdad puede considerarse como la definición de la proyección del rotor del campo dado sobre el vector unidad ν, arbitrariamente elegido pero fijo. Esto nos lleva también a la nueva definición del propio vector, puesto que será suficiente, por ejemplo, elegir arbitrariamente tres vectores ortogonales unidad ν_1, ν_2, ν_3, las proyecciones sobre los cuales, como es sabido, determinan univocamente cualquier vector.

Se puede mostrar que las magnitudes que figuran en el segundo miembro de la igualdad (52.27) no dependen de la elección del sistema de coordenadas, no obstante la concordancia de las orientaciones del vector ν y contorno Γ sí depende de la orientación del sistema de coordenadas: al pasar del sistema de coordenadas derecho al izquierdo, la concordancia de las orientaciones de ν y Γ según "la regla del sacacorchos" se sustituye por la concordancia según la regla inversa, de acuerdo con la cual, siendo fijada la orientación del vector ν, la del contorno Γ se cambia en opuesta. De este modo, la integral $\int\limits_\Gamma a\,dr$ cambia de signo, cuando varía la orientación del sistema de coordenadas, por lo cual, en virtud de la fórmula (52.27), cambia de signo también $\operatorname{rot} a$.

De lo dicho proviene que la fórmula de Stokes (52.22) es lícita no sólo para el sistema de coordenadas derecho, sino también para el izquierdo, puesto que, al variar

---

*) Como en el teorema 3 (según la "regla del sacacorchos").

**) Es evidente que las regiones citadas S siempre existen (¿por qué?).

***) Mediante $\operatorname{rot}_\nu a$ está designada la proyección del vector $\operatorname{rot} a$ sobre el vector ν, es decir, $\operatorname{rot}_\nu a = \operatorname{pr}_\nu \operatorname{rot} a$.

la orientación del sistema de coordenadas, tanto el miembro izquierdo, como el derecho de la igualdad (52.22) cambian de signo: siendo fijada la orientación del vector ν de la superficie S, cambian de signo tanto rot a como también el contorno Γ, si varía la orientación del sistema de coordenadas.

52.5. CAMPOS VECTORIALES SOLENOIDALES

Una región limitada para la cual se cumple el teorema de Ostrogradski — Gauss (véase el p. 52.3) se llamará en este punto *admisible*. Se denominará *admisible* la totalidad de superficies, si es la frontera de una región admisible.

Se ha observado en el p. 52.3 que el teorema de Ostrogradski — Gauss es válido para cualquiera región acotada cuya frontera se compone de un número finito de las superficies suaves a trozos. Por ello, cualquiera de tales regiones es admisible. Es justa, evidentemente, la afirmación recíproca: toda región admisible tiene frontera compuesta de un número finito de las superficies suaves a trozos, pues, de lo contrario no se podría ni siquiera hablar sobre las integrales de superficie a lo largo de la frontera.

El lector que prefiere utilizar tan sólo hechos demostrados, puede tomar en calidad de regiones y superficies admisibles aquellas, para las cuales fue demostrado en el presente curso el teorema de Ostrogradski — Gauss.

Definición 9. *Un campo vectorial* $a = a(x, y, z)$, *continuamente derivable en la región* G, *se denomina solenoidal en la misma, si el flujo del campo a través de la frontera orientada de cualquier región admisible* D, *cuya clausura* $\overline{D}$ *se dispone en* G: $\overline{D} \subset G$, *es igual a cero:*

$$\iint_{\partial D} a\,dS = 0. \tag{52.29}$$

La frontera ∂D de la región admisible D tiene dos orientaciones, generadas por las normales interior y exterior, respectivamente. Es evidente que si la condición (52.29) se cumple con una orientación, se cumplirá también con la otra, puesto que las integrales correspondientes pueden diferenciarse sólo en el signo.

Expliquemos la definición del campo solenoidal con un ejemplo. Sea G un anillo esférico: una parte del espacio comprendida entre dos esferas S_r y S_R que tienen el centro común O y los radios r y R, $r < R$, y sea a el campo vectorial solenoidal en G. Entonces, su flujo será nulo, por ejemplo, a través de cualquier esfera S que se dispone en G y limita una bola, también dispuesta en G.

Sin embargo, el flujo del campo vectorial a a través de la esfera S_ρ de radio ρ, $r < \rho < R$, y centro en el punto O no ha de ser nulo, puesto que la bola, limitada por dicha esfera, no se contiene en la región G (fig. 225).

Por otro lado, será nula la suma de los flujos del campo vectorial a través de dos esferas S_{ρ_1} y S_{ρ_2} de radios ρ_1 y ρ_2, $r < \rho_1 < \rho_2 < R$, y del mismo centro, si orientamos una de éstas eligiendo una normal que va al centro y la otra, en la dirección opuesta. Efectivamente, las esferas mencionadas limitan un anillo esférico, íntegramente dispuesto en la región G, mientras que su orientación elegida es la orientación de la frontera que corresponde a la normal exterior o interior. Por esto, según la definición del campo solenoidal, el flujo de éste a través de la frontera orientada que se considera será igual a cero.

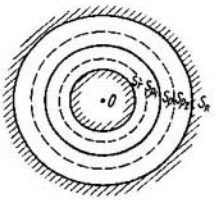


Fig. 225

Teorema 5. *Para que un campo vectorial continuamente derivable en la región G sea solenoidal en ella, es necesario y suficiente que su divergencia sea nula en todos los puntos de la región G:*

$$\operatorname{div} a(M) = 0, \quad M \in G.$$

DEMOSTRACIÓN DE LA NECESIDAD. Sea a un campo vectorial solenoidal en la región G y $M_0 \in G$. Designemos con Q_r una bola abierta de radio $r > 0$ y centro en el punto M_0, y con S_r, la esfera que limita dicha bola. Por cuanto todos los puntos $M \in G$, incluido también el punto M_0, son interiores para G, entonces existe tal $r_0 > 0$ que para $r < r_0$ todas las bolas de radio r, junto con las esferas que las limitan, se contendrán dentro de G.

Ha de notarse ahora que el límite (52.17), igual al valor de la divergencia del campo vectorial a en el punto M_0, existe para las regiones admisibles arbitrarias D, $D \subset \bar{D} \subset G$, cuyos diámetros tienden a cero. Por esta razón él existe también cuando $D = Q_r$, $r < r_0$ se elige especialmente:

$$\operatorname{div} a(M_0) = \lim_{r \to 0} \frac{\iint\limits_{S_r} a\,dS}{\mu Q_r}.$$

En vista de la definición del campo solenoidal, para todo $r < r_0$ tiene lugar la igualdad

$$\iint\limits_{S_r} a\,dS = 0,$$

por lo cual $\operatorname{div} a(M_0) = 0$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Sea a un campo vectorial, continuamente derivable en la región G, con la divergencia igual a cero en todos los puntos de la región G. Si D es una región admisible arbitraria de tal género que $D \subset \bar{D} \subset G$, entonces, en virtud del teorema de Ostrogradski — Gauss

$$\iint\limits_{\partial D} a\,dS = \iiint\limits_{D} \operatorname{div} a\,dx\,dy\,dz = 0,$$

es decir, el campo a es solenoidal. □

De ejemplo tipo de campo solenoidal interviene un campo vectorial que representa en sí en cierta región el campo de rotores de un campo vectorial dos veces continuamente derivable en la región citada.

En efecto, si a es un campo dos veces continuamente derivable en la región G, entonces rot a es un campo solenoidal en G, ya que div rot $a = 0$.

No es difícil ofrecer un razonamiento verosímil que afirma la validez de esta correlación. Con este fin basta pasar al vector simbólico ∇; entonces la igualdad en consideración tomará la forma $\nabla(\nabla \times a) = 0$.

Un producto mixto de los vectores ordinarios en el caso en que dos factores son coincidentes, es nulo, puesto que en este caso un paralelepípedo tendido sobre dichos vectores degenera en un paralelogramo y, por ende, su volumen es igual a cero. Por esta razón es natural esperar que la igualdad mencionada es verídica también para el vector ∇. Este razonamiento verosímil puede convertirse en el argumentado matemáticamente que tenga una fuerza probatoria, si demostramos que el vector simbólico ∇ posee, de hecho, las propiedades, ya utilizadas por nosotros, análogas a las propiedades correspondientes de los vectores ordinarios. Lo último puede hacerse con ayuda de una simple comprobación, pasando, por ejemplo, a una notación coordenada (véanse (52.2) y (52.4)).

52.6. CAMPOS VECTORIALES POTENCIALES

En este punto la superficie S, para la cual es válido el teorema de Stokes, se llamará admisible.

Definición 10. *La región tridimensional G se denomina simplemente conexa, si, cualquiera que sea una quebrada cerrada γ, dispuesta en G, existe la superficie admisible S que se dispone tambien en G y está tendida sobre la quebrada γ* (véase el punto 52.4).

A veces las regiones simplemente conexas se llaman *conexas simple y superficialmente*.

Si la región en consideración G es convexa, existe un método muy simple para tender las superficies sobre el contorno. La superficie buscada siempre puede tomarse en este caso en forma de un cono con vértice en el punto arbitrariamente fijado $M_0 \in G$, para el cual la curva dada γ sirve como directriz. Si

$$\rho = \rho(u), \quad 0 \leqslant u \leqslant 2\pi,$$

es la representación de esta curva y r_0 es el radio vector del punto M_0, entonces el cono buscado S, tendido sobre el contorno dado, se da mediante la representación

$$r = r_0 + v[\rho(u) - r_0], \quad 0 \leqslant u \leqslant 2\pi, \quad 0 \leqslant v \leqslant 1. \tag{52.30}$$

Considerando u y v como coordenadas polares, vemos que la "representación" del cono viene dada en un círculo unitario, con la particularidad de que la circunferencia unitaria $\gamma_0 = \{(u, v) : v = 1\}$ pasa al contorno dado γ, y su centro, al vértice del cono (fig. 226).

La palabra "representación" viene entre comillas, puesto que el concepto de representación de una superficie se ha introducido más arriba sólo para el caso en que los parámetros u y v eran coordenadas cartesianas. El cono (52.30) tendrá, en el

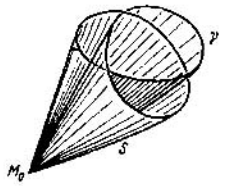


Fig. 226

caso general, puntos múltiples y no será una superficie suave a trozos incluso en el caso cuando γ sea una curva suficientemente suave, es decir, si γ es una curva sin puntos singulares continuamente derivable un número suficiente de veces. En el cono (52.30) habrá, en general, puntos singulares distintos del vértice. Precisamente para eliminar este inconveniente de un modo más sencillo, nos hemos limitado, definiendo la región simplemente conexa, a considerar los contornos en forma de las quebradas cerradas. En este último caso el vértice del cono M_0 puede siempre elegirse de un modo tal que el cono citado sea una superficie suave a trozos. Efectivamente, cualquiera que sea la elección del vértice del cono en el caso en que una cierta quebrada γ sirva de su directriz, el cono se descompone en un número finito de triángulos S_i, $i = 1, 2, \ldots, k$, los cuales, en verdad, serán, quizás, degenerados, es decir se convertirán en un segmento o un punto. Uno de los vértices de dichos triángulos será precisamente el vértice M_0 del cono, mientras que el lado opuesto lo constituirá uno de los lados de la quebrada γ. Cada uno de estos triángulos puede considerarse como una superficie continuamente derivable cualquier número de veces y prefijarse mediante una representación realizada por las funciones lineales (véanse el p. 16.5 y (52.30)). Si el triángulo es degenerado, todos los puntos de él serán singulares. No obstante, mediante un desplazamiento, tan pequeño como se quiera, del vértice del cono podemos conseguir que el vértice esté en una posición general con todos los lados de la quebrada γ, es decir, no se disponga en ninguna de las rectas que pasé por un lado cualquiera de la quebrada γ. De resultas, todos los triángulos S_i, $i = 1, 2, \ldots, k$, serán regulares y, por lo tanto, podrán considerarse como superficies suaves privadas de puntos singulares. En cuanto al propio cono S, éste resultará ser, de tal modo, una superficie suave a trozos $S = \{S_i\}_{i=1}^{i=k}$. Además, por cuanto todo punto de una región, cualquiera que sea su desplazamiento tan pequeño como se quiera, queda dentro de la región, el vértice M_0 del cono S puede siempre elegirse en la misma, a consecuencia de lo cual, por ser convexa la región, todo el cono se dispondrá dentro de ella. Al cono suave a trozos S, que se ha obtenido, sé le puede aplicar el teorema de Stokes, en otras palabras, dicho cono es una superficie admisible en este punto. Hemos demostrado pues que toda *región convexa es* simplemente *conexa*.

Como ejemplo de región no simplemente conexa puede servir un toro, es decir, una región formada por la rotación de un círculo en torno del eje que no lo interseca (fig. 227).

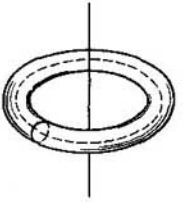


Fig. 227

Recordemos que un campo se llama potencial, si su circulación $\int_\gamma a\,dr$ es nula a lo largo de cualquier contorno cerrado $\gamma \subset G$, o bien, que es lo mismo, si la integral $\int_{\overset{\frown}{AB}} a\,dr$ no depende del camino de integración que une en la región G los puntos A y B. Los razonamientos más detallados sobre esto se encuentran en el p. 52.1. Resulta que en una región simplemente conexa un campo vectorial será potencial, si, y sólo si, es irrotacional. Esta afirmación está contenida y se demuestra en el teorema 6 que sigue abajo.

Teorema 6. *Sea dado en la región simplemente conexa G un campo vectorial continuamente derivable $a = (P, Q, R)$. Para este caso son equivalentes las siguientes tres propiedades:*

1. *El campo vectorial $a = a(M)$ es potencial en la región G.*

2. *Existe una función $u = u(M)$, potencial en G, es decir, tal función $u(M)$ que $a = \operatorname{grad} u$, o bien, que es lo mismo, $du = P\,dx + Q\,dy + R\,dz$. En este caso para cualesquiera dos puntos $A \in G$ y $B \in G$ y toda curva suave a trozos $\overset{\frown}{AB}$, que en G une dichos puntos, se verifica la igualdad*

$$\int_{\overset{\frown}{AB}} a\,dr = u(B) - u(A).$$

3. *El campo vectorial $a = a(M)$ es irrotacional:* $\operatorname{rot} a = 0$ *en la región G, es decir,*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x},\quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y},\quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}.$$

Subrayemos que del teorema 6 se deduce, en particular, que *el campo vectorial a, continuamente derivable en una región simplemente conexa, es potencial cuando, y sólo cuando, sea un campo de gradientes de cierta función escalar u:*

$$a = \nabla u.$$

DEMOSTRACIÓN. Se realizarán los razonamientos según el siguiente esquema.

Paso primero. $1 \Rightarrow 2$. Esta afirmación, es decir, la existencia de la función potencial, se demuestra por analogía completa con el caso de una región plana considerado anteriormente (véase el teorema 3 en el p. 47.8) y por esta razón nos abstenemos de su demostración.

Paso segundo: $2 \Rightarrow 3$. La afirmación $2 \Rightarrow 3$ también se demuestra igual que en el caso plano: significa simplemente que las correspondientes segundas derivadas mixtas de una función potencial son iguales.

Las afirmaciones $1 \Rightarrow 2$ y $2 \Rightarrow 3$ son justas también sin que se suponga que G es una región simplemente conexa.

Paso tercero: $3 \Rightarrow 1$. Sea en G $\operatorname{rot} a = 0$. Admitamos primero que γ es una curva cerrada, dos veces continuamente derivable a trozos, dispuesta en G. Si existe una superficie admisible S, contenida en G y limitada por el contorno γ, entonces, del teorema de Stokes obtenemos inmediatamente

$$\int_{\gamma} a\,dr = \iint_{S} \operatorname{rot} a\,dS = 0.$$

Ya que la región G es simplemente conexa (véase la definición 10), esta igualdad se verifica, en particular, para cualquier quebrada que tiene un número finito de lados. Por eso, si γ es una curva cerrada suave a trozos, dispuesta en G, entonces, eligiendo una sucesión de quebradas λ_n, inscritas en γ con los lados que tienden a cero para $n \to \infty$, de acuerdo con el lema 3 del p. 47.8, obtendremos

$$\int_{\gamma} a\,dr = \lim_{n \to \infty} \int_{\gamma_n} a\,dr = 0. \quad \square$$

Como conclusión observemos que aunque los campos vectoriales potenciales y solenoidales no abarcan todos los campos vectoriales posibles, permiten, sin embargo, describir una clase bastante amplia de los campos vectoriales. A saber, asumidas ciertas suposiciones suficientemente generales, cualquier campo vectorial a representa una suma de los campos vectoriales potencial y solenoidal. Con más precisión, existen una función escalar u y un campo vectorial b tales que $a = \nabla u + \operatorname{rot} b$. Por cuanto $\operatorname{rot} \nabla u = 0$ y $\operatorname{div} \operatorname{rot} b = 0$, entonces el primer sumando es un campo potencial y el segundo, solenoidal.

Esta suposición se llama también teorema de Helmholtz *) (su demostración se puede hallar en el libro "Métodos de la física teórica" que se debe a la pluma de P. M. Morse y G. Feshbach (P. Morse, G. Feshbach, *Methods of theoretical physics*, New York-Toronto-London, 1953).

Ejercicios. 20. Demuéstrese que el flujo de un rotor de un campo vectorial continuamente derivable en cierta región a través de una esfera cualquiera dispuesta en la región mencionada es igual a cero.

21. Demuéstrese que

$$\iiint_{G} \operatorname{grad} \varphi \operatorname{rot} a\,dx\,dy\,dz = \iint_{S} (a \times \operatorname{grad} \varphi\,dS).$$

*) H. Helmholtz (1821—1894), físico y fisiólogo alemán.

Aquí se supone que para la región G, limitada por la superficie S, queda aplicable el teorema de Ostrogradski — Gauss.

22. Hállense la divergencia y el rotor para los campos vectoriales $a = r/|r|^3$ y $b = r/|r|$. ¿Serán estos campos potenciales, solenoidales? Calcúlese el flujo de los campos vectoriales a y b a través de esfera $S_R = \{(x, y, z) : x^2 + y^2 + z^2 = R^2\}$.

23. Calcúlese el flujo del campo vectorial $a = \dfrac{r}{|r|^3}$ a través de la esfera $x^2 + y^2 + (z - 1)^2 = 2$.

24. Sean a, b y c los campos vectoriales derivables y sea u una función escalar, dos veces derivable, en la región $G \subset R^3$, $b = \operatorname{grad} u$, $a = b + c$. Demuéstrese que para que sea $\operatorname{div} c = 0$, es necesario y suficiente que la función u satisfaga en G la ecuación $\Delta u = \operatorname{div} a$ (de este modo la demostración del teorema de Helmholtz se reduce a la resolución en la región G de la ecuación del tipo $\Delta u = f(x, y, z)$).

Haciendo uso del teorema de Ostrogradski — Gauss, calcúlese el flujo del campo vectorial a a través de la superficie cerrada S, si:

25. $a = (1 + 2x)i + yj + zk$, $S = \{(x, y, z) : x^2 + y^2 = z^2, 0 \leqslant z \leqslant 4\}$.

26. $a = 2xi - yj + zk$, $S = \{(x, y, z) : x^2 + y^2 + z^2 = 4, 3z = x^2 + y^2\}$.

Establézcase cuáles de los siguientes campos vectoriales son solenoidales:

27. $a = x(z^2 - y^2)i + y(x^2 - z^2)j + z(y^2 + x^2)k$.

28. $a = (1 + 2xy)i - y^2zj + (z^2y - 2yz + 1)k$.

Haciendo uso del teorema de Stokes, hállese la circulación del vector a a lo largo del contorno γ, si

29. $a = yi - xj + zk$; $\gamma = \{(x, y, z) : x^2 + y^2 + z^2 = 4, x^2 + y^2 = z^2, z \geqslant 0\}$.

30. $a = y^2i + z^2j$; $\gamma = \{(x, y, z) : x^2 + y^2 = 9, 3y + 4z = 5\}$.

§ 53. INTEGRALES PROPIAS DEPENDIENTES DE UN PARÁMETRO

53.1. DEFINICIÓN DE LAS INTEGRALES DEPENDIENTES DE UN PARÁMETRO; SU CONTINUIDAD E INTEGRABILIDAD SEGÚN EL PARÁMETRO

Supongamos que Y es un conjunto de números reales, $\varphi(y)$ y $\psi(y)$ son dos funciones definidas en Y, $\varphi(y) \leqslant \psi(y)$ y la función $f(x, y)$ está definida en el conjunto

$$\{(x, y) : y \in Y, x \in [\varphi(y), \psi(y)]\}. \tag{53.1}$$

Las integrales de la forma

$$\Phi(y) = \int_{\varphi(y)}^{\psi(y)} f(x, y)dx \tag{53.2}$$

se denominan *integrales dependientes de un parámetro*, y la variable y se llama, corrientemente, *parámetro*.

Se encuentra a menudo el caso particular de tal tipo de integrales, cuando las funciones φ y ψ son constantes, es decir, las integrales de la forma

$$\Phi(y) = \int_a^b f(x, y)dx. \tag{53.3}$$

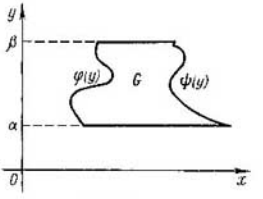


Fig. 228

Si Y es un conjunto de todos los números naturales, $Y = N = \{1, 2, \ldots, n, \ldots\}$, entonces, suponiendo $f(x, n) = f_n(x)$, $n = 1, 2, \ldots$, la integral (53.3) puede escribirse en la forma

$$\int_a^b f_n(x)dx, \quad n = 1, 2, \ldots .$$

Se ha obtenido de este modo una sucesión numérica formada por las integrales de las funciones de cierta sucesión funcional.

Examinemos un caso en que el conjunto Y representa en sí un segmento $[\alpha, \beta]$, las funciones $\varphi(y)$ y $\psi(y)$ son continuas en dicho segmento y $\varphi(y) \leqslant \psi(y)$, $y \in [\alpha, \beta]$. Supongamos que las gráficas de las funciones $\varphi(y)$ y $\psi(y)$ y, quizás, los segmentos de las rectas $y = \alpha$ e $y = \beta$ forman la frontera de la región limitada G (fig. 228). Esta región es, evidentemente, cuadrable (véase el p. 44.1). En este caso el conjunto (53.1), sobre el cual viene definida la función $f(x, y)$, es la clausura $\bar{G}$ de la región mencionada G:

$$\bar{G} = \{(x, y); \alpha \leqslant y \leqslant \beta, \varphi(y) \leqslant x \leqslant \psi(y)\}. \tag{53.4}$$

En lo sucesivo estudiaremos las propiedades de la función $\Phi(y)$ (su continuidad, las reglas de su derivación e integración) en dependencia de las propiedades de las funciones $f(x, y)$, $\varphi(y)$, $\psi(y)$. Algunas de dichas propiedades se han obtenido antes, al estudiar la integral múltiple. Así por ejemplo, el lema demostrado en el p. 45.1 ofrece las condiciones, bajo las cuales una integral dependiente de un parámetro es una función continua de este parámetro. Enunciemos este lema en forma de un teorema, aplicando las designaciones del presente párrafo.

Teorema 1. *Si la función $f(x, y)$ es continua en la clausura $\bar{G}$ de la región G* (véase (53.4)), *la función $\Phi(y)$, definida por la fórmula* (53.2), *es continua en el segmento* $[\alpha, \beta]$.

A la afirmación de este teorema se le puede comunicar la siguiente expresión:

$$\lim_{y \to y_0} \int_{\varphi(y)}^{\psi(y)} f(x, y)dx = \int_{\lim\limits_{y \to y_0} \varphi(y)}^{\lim\limits_{y \to y_0} \psi(y)} \lim_{y \to y_0} f(x, y)dx. \tag{53.5}$$

En efecto, del teorema 1 se deduce que el límite en el primer miembro de la igualdad (53.5) es igual a $\Phi(y_0)$, y, en virtud de la continuidad de las funciones φ, ψ y f, el segundo miembro de la igualdad es también igual a

$$\int_{\varphi(y_0)}^{\psi(y_0)} f(x, y_0)dx = \Phi(y_0).$$

En particular, para la integral (53.3) tenemos

$$\lim_{y \to y_0} \int_a^b f(x, y)dx = \int_a^b \lim_{y \to y_0} f(x, y)dx,$$

es decir, en este caso resulta posible el paso límite bajo el signo de integral.

En el teorema sobre el paso límite bajo el signo de integral se pueden debilitar las exigencias impuestas sobre la función $f(x, y)$, al requerir solamente, en lugar de su continuidad respecto de la totalidad de variables, que sea continua respecto de una sola variable y que tienda uniformemente al límite, respecto de la otra.

Teorema 2. *Supongamos que la función $f(x, y)$ está definida para todo $x \in [a, b]$, $y \in Y$ y es continua respecto de x en $[a, b]$ para cualquier y fijo, $y \in Y$. Entonces, si para $y \to y_0$ *) la función $f(x, y)$ tiende uniformemente en el segmento $[a, b]$ hacia la función $\varphi(x)$* (véase el p. 39.4), *se tiene*

$$\lim_{y \to y_0} \int_a^b f(x, y)dx = \int_a^b \varphi(x)dx.$$

DEMOSTRACIÓN. Consideremos una sucesión $y_n \in Y$, $n = 1, 2, \ldots$, tal que $\lim_{y \to y_0} y_n = y_0$. Entonces (véase el ejercicio 5 en el p. 39.4) la sucesión $\varphi_n(x) = f(x, y_n)$ tenderá uniformemente en el segmento $[a, b]$ hacia la función $\varphi(x)$. De aquí se deduce (véase el p. 36.4) en primer lugar que $\varphi(x)$ es continua y, por ende, integrable en el segmento $[a, b]$, y, en el segundo lugar, que

$$\lim_{n \to \infty} \int_a^b f(x, y_n)dx = \lim_{n \to \infty} \int_a^b \varphi_n(x)dx = \int_a^b \varphi(x)dx,$$

y, como esto es cierto para cualquier sucesión indicada $\{y_n\}$, el teorema queda demostrado. □

Pasemos al problema de integrabilidad de las integrales (53.2) que dependen de un parámetro.

Teorema 3. *Supongamos que la región G es elemental respecto de ambos ejes de coordenadas, es decir,*

$$G = \{(x, y) : \alpha < y < \beta, \varphi(y) < x < \psi(y)\} =$$
$$= \{(x, y) : a < x < b, \varphi_1(x) < y < \psi_1(x)\},$$

*) Aquí y_0 es un número o uno de los infinitos: ∞, $+\infty$ ó $-\infty$.

donde las funciones φ y ψ son continuas en el segmento $[\alpha, \beta]$, mientras que las funciones φ_1 y ψ_1 lo son en el segmento $[a, b]$. Entonces, si la función $f(x, y)$ es continua en la clausura $\bar{G}$ de la región G, se tiene

$$\int_\alpha^\beta \Phi(y)dy = \int_\alpha^\beta \left[\int_{\varphi(y)}^{\psi(y)} f(x, y)dx\right]dy = \int_a^b \left[\int_{\varphi_1(x)}^{\psi_1(x)} f(x, y)dy\right]dx =$$

$$= \iint_G f(x, y)dxdy. \quad (53.6)$$

Es evidente que el teorema 3 hace paráfrasis del teorema correspondiente sobre la reducción de la integral múltiple a una reiterada (véase el p. 45.1).

53.2. DERIVACIÓN DE LAS INTEGRALES DEPENDIENTES DE UN PARÁMETRO

Al estudiar las propiedades diferenciales de las integrales que dependen de un parámetro, consideraremos al principio las integrales de la forma (53.3).

Teorema 4 (regla de Leibniz). *Si la función $f(x, y)$ y su derivada parcial $\dfrac{\partial f(x, y)}{\partial y}$ son continuas en un rectángulo cerrado $P = \{(x, y) : a \leqslant x \leqslant b, \alpha \leqslant y \leqslant \beta\}$, entonces la función $\Phi(y) = \int_a^b f(x, y)dx$ es derivable en el segmento $[\alpha, \beta]$ y*

$$\frac{d\Phi(y)}{dy} = \int_a^b \frac{\partial f(x, y)}{\partial y}dx.$$

Así pues, para derivar, bajo las suposiciones asumidas, una integral que depende de parámetro, es suficiente derivar la expresión subintegral, dejando intactos los límites de integración.

DEMOSTRACIÓN. Sea $y \in [\alpha, \beta]$ e $y + \Delta y[\alpha, \beta]$; entonces

$$\frac{\Phi(y + \Delta y) - \Phi(y)}{\Delta y} = \frac{1}{\Delta y}\int_a^b [f(x, y + \Delta y) - f(x, y)]dx =$$

$$= \int_a^b \frac{\partial f(x, y + \theta\Delta y)}{\partial y}dx, \quad 0 < \theta < 1.$$

Se ha aplicado aquí la fórmula de incrementos finitos de Lagrange.

Al designar ahora con $\omega\left(\delta; \dfrac{\partial f}{\partial y}\right)$ el módulo de continuidad de la función $\dfrac{\partial f}{\partial y}$, obtendremos

$$\left|\frac{\Phi(y + \Delta y) - \Phi(y)}{\Delta y} - \int_a^b \frac{\partial f(x, y)}{\partial y}dx\right| \leqslant$$

$$\leqslant \int_a^b \left|\frac{\partial f(x, y + \theta\Delta y)}{\partial y} - \frac{\partial f(x, y)}{\partial y}\right|dx \leqslant \int_a^b \omega\left(|\Delta y|; \frac{\partial f}{\partial y}\right)dx \leqslant$$

$$\leqslant \omega\left(|\Delta y|; \frac{\partial f}{\partial y}\right)(b-a). \quad (53.7)$$

Dado que la función $\dfrac{\partial f}{\partial y}$ es uniformemente continua en el rectángulo cerrado P, resulta que $\lim\limits_{\Delta y \to 0} \omega\left(|\Delta y|; \dfrac{\partial f}{\partial y}\right) = 0$; por eso, de (53.7) obtenemos

$$\lim_{\Delta y \to 0} \frac{\Phi(y+\Delta y)-\Phi(y)}{\Delta y} = \int_a^b \frac{\partial f(x,y)}{\partial y}\,dx. \square$$

El teorema 4 se generaliza con facilidad para el caso de una integral de la forma general (53.2) que depende de un parámetro.

Teorema 4′. *Supongamos que:*

1) *la función $f(x, y)$ y su derivada parcial $\dfrac{\partial f(x,y)}{\partial y}$ son continuas en el rectángulo cerrado*

$$P = \{(x, y) : a \leqslant x \leqslant b, \alpha \leqslant y \leqslant \beta\},$$

2) $\bar{G} \subset P$ (véase (53.4));

3) *las funciones $\varphi(y)$ y $\psi(y)$ tienen derivadas continuas en el segmento $[\alpha, \beta]$.*

Bajo estas condiciones la integral (53.2), *dependiente de un parámetro, tiene también una derivada en el segmento $[\alpha, \beta]$, con la particularidad de que*

$$\frac{d\Phi}{dy} = \int_{\varphi(y)}^{\psi(y)} \frac{\partial f(x,y)}{\partial y}\,dx - f[\varphi(y), y]\frac{d\varphi(y)}{dy} + f[\psi(y), y]\frac{d\psi(y)}{dy}. \quad (53.8)$$

DEMOSTRACIÓN. Examinemos una función

$$F(y, u, v) = \int_u^v f(x, y)dx, \quad a \leqslant u \leqslant b, \quad a \leqslant v' \leqslant b, \quad \alpha \leqslant y \leqslant \beta.$$

No es difícil comprobar directamente que las derivadas parciales $\dfrac{\partial F}{\partial y}, \dfrac{\partial F}{\partial u}, \dfrac{\partial F}{\partial v}$ de la función F existen y son continuas en totalidad de las variables y, u, v. Comprobemos primero la existencia y continuidad de la derivada parcial $\dfrac{\partial F}{\partial y}$. Su existencia se desprende inmediatamente del teorema 4, con la particularidad de que

$$\frac{\partial F}{\partial y} = \int_u^v \frac{\partial f(x,y)}{\partial y}dx. \quad (53.9)$$

Demostremos su continuidad. Sea $a \leqslant u \leqslant b$, $a \leqslant v \leqslant b$, $\alpha \leqslant y \leqslant \beta$, $a \leqslant u + \Delta u \leqslant b$, $a \leqslant v + \Delta v \leqslant b$, $\alpha \leqslant y + \Delta y \leqslant \beta$; al poner

$$\Delta \frac{\partial F(y, u, v)}{\partial y} = \frac{\partial F(y+\Delta y, u+\Delta u, v+\Delta v)}{\partial y} - \frac{\partial F(y, u, v)}{\partial y},$$

obtendremos:

$$\left|\Delta\frac{\partial F(y,u,v)}{\partial y}\right| = \left|\int_{u+\Delta u}^{v+\Delta v}\frac{\partial f(x,y+\Delta y)}{\partial y}dx - \int_{u}^{v}\frac{\partial f(x,y)}{\partial y}dx\right| \leqslant$$

$$\leqslant \left|\int_{u}^{v}\left[\frac{\partial f(x,y+\Delta y)}{\partial y} - \frac{\partial f(x,y)}{\partial y}\right]dx\right| +$$

$$+ \left|\int_{u+\Delta u}^{u}\frac{\partial f(x,y+\Delta y)}{\partial y}dx\right| + \left|\int_{v}^{v+\Delta v}\frac{\partial f(x,y+\Delta y)}{\partial y}dx\right|. \quad (53.10)$$

Por cuanto la función $\dfrac{\partial f}{\partial y}$ está definida en el rectángulo P, todas las integrales escritas tienen sentido en virtud de la elección indicada anteriormente de valores de los argumentos y

$$|v-u| \leqslant b-a. \quad (53.11)$$

Luego, de la continuidad de la función $\dfrac{\partial f}{\partial y}$ en el rectángulo P proviene que esta función está acotada en él, es decir, existe tal constante $M > 0$ que para todo punto $(x, y) \in P$ se verifica la desigualdad

$$\left|\frac{\partial f(x,y)}{\partial y}\right| \leqslant M. \quad (53.12)$$

Al designar, como hasta ahora, con $\omega\left(\delta; \dfrac{\partial f}{\partial y}\right)$ el módulo de continuidad de la función $\dfrac{\partial f}{\partial y}$ en el rectángulo P y hacer uso de las desigualdades (53.11) y (53.12), de (53.10) obtenemos

$$\left|\Delta\frac{\partial F(y,u,v)}{\partial y}\right| \leqslant \omega\left(|\Delta y|; \frac{\partial f}{\partial y}\right)\left|\int_{u}^{v}dx\right| + M\left|\int_{u+\Delta u}^{u}dx\right| +$$

$$+ M\left|\int_{v}^{v+\Delta v}dx\right| \leqslant (b-a)\omega\left(|\Delta y|; \frac{\partial f}{\partial y}\right) + M|\Delta u| + M|\Delta v|.$$

De aquí se deduce que $\lim\limits_{\sqrt{\Delta y^2+\Delta u^2+\Delta v^2}\to 0}\Delta\dfrac{\partial F(y,u,v)}{\partial y} = 0$. Esto significa precisamente la continuidad de la derivada parcial $\dfrac{\partial F}{\partial y}$ en el conjunto

$$\{(y, u, v) : c \leqslant y \leqslant d, a \leqslant u \leqslant b, a \leqslant v \leqslant b\}.$$

La continuidad en este conjunto de las derivadas parciales

$$\frac{\partial F}{\partial u} = -f(u,v), \quad \frac{\partial F}{\partial v} = f(v,y) \quad (53.13)$$

es obvia.

La relación entre las funciones Φ y F se establece por medio de la fórmula

$$\Phi(y) = F(y, \varphi(y), \psi(y)).$$

Debido a lo demostrado anteriormente, la función Φ puede derivarse según la regla para derivar las funciones compuestas:

$$\frac{\partial \Phi}{\partial y} = \frac{\partial F}{\partial y} + \frac{\partial F}{\partial u}\frac{du}{dy} + \frac{\partial F}{\partial v}\frac{dv}{dy}.$$

Sustituyendo aquí las expresiones para las derivadas parciales

$$\frac{\partial F}{\partial y}, \frac{\partial F}{\partial u} \text{ y } \frac{\partial F}{\partial v}$$

(véase (53.9) y (53.13)) y suponiendo $u = \varphi(y)$ y $v = \psi(y)$, obtendremos la fórmula (53.8). □

§ 54. INTEGRALES IMPROPIAS DEPENDIENTES DE UN PARÁMETRO

54.1. DEFINICIONES FUNDAMENTALES. CONVERGENCIA UNIFORME DE LAS INTEGRALES DEPENDIENTES DE UN PARÁMETRO

Consideraremos las integrales del tipo

$$\Phi(y) = \int_a^b f(x, y)dx, \tag{54.1}$$

donde $-\infty \leqslant a < b \leqslant +\infty$, la variable y pertenece a cierto conjunto Y y la integral (54.1) es, para ciertos valores de y (en particular para todo y), impropia.

Definición 1. *Si, para todo $y_0 \in Y$, la integral*

$$\Phi(y_0) = \int_a^b f(x, y_0)dx$$

converge, la integral (54.1) *se llama convergente en el conjunto Y.*

En lo sucesivo, si no se especifica alguna otra circunstancia, consideraremos sólo el caso en que se cumplen las condiciones:

1) $-\infty < a < b \leqslant +\infty$;

2) cualquiera que sea $y \in Y$, la función $f(x, y)$ es integrable según Riemann respecto de la variable x en todo segmento $[a, \eta]$, donde η es tal que $a < \eta < b$.

En este caso la convergencia de la integral (54.1) en el conjunto Y significa que para cualquier $y \in Y$ existe un límite

$$\lim_{\eta \to b-0} \int_a^\eta f(x, y)dx = \int_a^b f(x, y)dx$$

(si es que $b = +\infty$, entonces $b - 0 = +\infty$). Por cuanto

$$\int_a^b f(x, y)dx - \int_a^\eta f(x, y)dx = \int_\eta^b f(x, y)dx,$$

de lo dicho obtendremos, para todo $y \in Y$ fijo:

$$\lim_{\eta \to b-0} \int_{\eta}^{b} f(x, y)dx = 0.$$

De este modo, *si la integral* (54.1) *converge en el conjunto* Y, *entonces para todo* $y \in Y$ *fijo, cualquiera que sea el número* $\varepsilon > 0$, *existe tal* $\eta_\varepsilon = \eta_\varepsilon(y) < b$, *que si* $\eta_\varepsilon \leqslant \eta < b$, *se tiene*

$$\left| \int_{\eta}^{b} f(x, y)dx \right| < \varepsilon. \tag{54.2}$$

Las condiciones, bajo las cuales para las integrales impropias dependientes de un parámetro resultan válidos los teoremas, análogos a los demostrados en el párrafo anterior para integrales propias, están basadas sobre el concepto de la así llamada convergencia uniforme de la integral.

Se supondrá, según lo observado más arriba, que la integral (54.1) satisface las condiciones citadas 1) y 2).

Definición 2. *La integral* $\int_{a}^{b} f(x, y)dx$, *convergente en el conjunto* Y, *se llama convergente uniformemente en dicho conjunto, si para cualquier* $\varepsilon > 0$ *existe tal* $\eta_\varepsilon < b$ *que para todo* $y \in Y$ *y todos los* η *tales que* $\eta_\varepsilon < \eta < b$, *se verifica la desigualdad*

$$\left| \int_{\eta}^{b} f(x, y)dx \right| < \varepsilon.$$

Recordemos que en nuestro caso b puede ser tanto finito, o sea, un número, como infinito, o sea, igual a $+\infty$. De este modo, la definición de convergencia uniforme en la forma aducida es válida simultáneamente tanto para el caso en que la integración se realiza por el segmento finito $[a, b]$, mientras que la integral impropia surja a cuenta de que el integrando no está acotado, como para el caso en que la integral impropia aparece a cuenta de que no está acotado el intervalo de integración $[a, +\infty]$.

Las definiciones aducidas de la convergencia y convergencia uniforme de una integral recuerdan las definiciones correspondientes para las series (véanse los pp. 36.1 y 36.3). Entre estas definiciones realmente existe una cierta relación.

Sea $\{\eta_n\}$ una sucesión tal que

$$\eta_1 = a, \eta_n \in [a, b), \quad n = 1, 2, \ldots, \quad \text{y} \quad \lim_{n \to \infty} \eta_n = b.$$

A la par con la integral (54.1) consideraremos la serie

$$\sum_{k=1}^{\infty} \int_{\eta_k}^{\eta_{k+1}} f(x, y)dx. \tag{54.3}$$

Sea

$$S_n(y) = \sum_{k=1}^{n-1} \int_{\eta_k}^{\eta_{k+1}} f(x, y)dx = \int_{a}^{\eta_n} f(x, y)dx$$

su suma parcial. En este caso, si la integral (54.1) converge (converge uniformemente) sobre el conjunto Y, será, evidentemente, convergente (uniformemente convergente) sobre el conjunto Y también la serie (54.3); con ello

$$\int_a^b f(x,y)dx = \lim_{n\to\infty} \int_a^{\eta_n} f(x,y)dx = \lim_{n\to\infty} S_n(y),$$

es decir, la integral en consideración es igual a la suma de la serie (54.3).

La definición de la convergencia uniforme puede parafrasearse además de la manera siguiente.

Definición 2'. *La integral* (54.1), *convergente sobre el conjunto* Y, *se llama uniformemente convergente sobre dicho conjunto*, *si*

$$\lim_{\eta\to b-0} \sup_{y\in Y} \left| \int_\eta^b f(x,y)dx \right| = 0. \tag{54.4}$$

En efecto, si la integral (54.1) converge uniformemente sobre el conjunto Y en el sentido de la definición 2, entonces para todo $\varepsilon > 0$ existe tal $\eta_\varepsilon < b$ que se cumple la desigualdad (54.2) siendo $y \in Y$ e $\eta_\varepsilon \leqslant \eta < b$, y por consiguiente,

$$\sup_{y\in Y} \left| \int_\eta^b f(x,y)dx \right| \leqslant \varepsilon, \eta_\varepsilon \leqslant \eta < b,$$

de donde proviene precisamente (54.4). Viceversa, si la integral en consideración converge uniformemente sobre el conjunto Y en el sentido de la definición 2', entonces de la condición (54.4) se infiere, para cualquier $\varepsilon > 0$, la existencia de tal número η_ε que con $y \in Y$ y $\eta_\varepsilon \leqslant \eta < b$ se verifica la desigualdad (54.2). □

Si consideramos la integral

$$F(y,\eta) \stackrel{\text{def}}{=} \int_\eta^b f(x,y)dx,$$

llegaremos a que, evidentemente, la condición (54.4) implica que esta integral tiende uniformemente sobre Y hacia cero para $\eta - b - 0$ (en la terminología del p. 39.4, el parámetro aquí no es y, como lo fue en el punto mencionado, sino la variable η).

La convergencia uniforme sobre el conjunto Y de la integral (54.1) significa también que en este conjunto la función

$$\Phi(y,\eta) \stackrel{\text{def}}{=} \int_a^\eta f(x,y)dx \tag{54.5}$$

tiende uniformemente, para $\eta - b - 0$, hacia la función (54.1).

Efectivamente, lo último significa (véase el p. 39.4) que para todo $\varepsilon > 0$ existe tal $\eta_\varepsilon < b$, que para cualquier η que satisface la condición $\eta_\varepsilon \leqslant \eta < b$, y cualquier $y \in Y$ se verifica la desigualdad

$$|\Phi(y) - \Phi(y,\eta)| < \varepsilon.$$

Pero

$$\Phi(y) - \Phi(y,\eta) = \int_a^b f(x,y)dx - \int_a^\eta f(x,y)dx = \int_\eta^b f(x,y)dx.$$

Por eso

$$\left|\int_{\eta}^{b} f(x,y)dx\right| < \varepsilon.$$

De este modo, la condición $\Phi(y,\eta) \rightrightarrows_{Y} \Phi(y)$ para $\eta \to b-0$ es equivalente al cumplimiento de las condiciones de la definicion 2, es decir, a la convergencia uniforme sobre el conjunto Y de la integral (54.1).

Ejemplo. Consideremos una integral $\Phi(y) = \int_{0}^{+\infty} ye^{-xy}dx$. A título de conjunto Y tomemos el semieje $y \geqslant 0$ (para cualquier $y < 0$ esta integral diverge). Es fácil convencerse de que la integral en consideración converge sobre Y. Para cualquier $\alpha > 0$ converge uniformemente en el intervalo $[\alpha, +\infty)$. Efectivamente, en este caso se comprueba con facilidad, por ejemplo, el cumplimiendo de la condición (54.4):

$$\lim_{\eta \to +\infty} \sup_{y \geqslant \alpha} \left|\int_{\eta}^{+\infty} ye^{-xy}dx\right| = \lim_{\eta \to +\infty} \sup_{y \geqslant \alpha} e^{-\eta y} = \lim_{\eta \to +\infty} e^{-\alpha\eta} = 0.$$

Mientras tanto, la convergencia uniforme no subsiste en todo el semieje. En efecto,

$$\lim_{\eta \to +\infty} \sup_{y \geqslant 0} \left|\int_{\eta}^{+\infty} ye^{-xy}dx\right| = \lim_{\eta \to +\infty} \sup_{y \geqslant 0} e^{-\eta y} = 1,$$

es decir, en el conjunto Y la condición (54.4) no se cumple.

Teorema 1 (criterio de Weierstrass). *Si existe una función no negativa $\varphi(x)$, definida en el intervalo $[a, b)$ e integrable según Riemann en todo segmento $[a, \eta]$, donde $a < \eta < b$, de tal índole que:*

1) $|f(x,y)| \leqslant \varphi(x)$, *donde* $a \leqslant x < b$, $y \in Y$;

2) *la integral* $\int_{a}^{b} \varphi(x)dx$ *converge*,

entonces la integral (54.1) *converge uniformemente sobre el conjunto* Y.

DEMOSTRACIÓN. Ante todo, en virtud del criterio de comparación (véase el p. 33.3), la integral (54.1) converge absolutamente y, por lo tanto, es simplemente convergente, cualquiera que sea $y \in Y$. Luego, por lo que la integral $\int_{a}^{b} \varphi(x)dx$ es convergente, para todo $\varepsilon > 0$ existe tal $\eta_{\varepsilon} < b$, que si $\eta_{\varepsilon} \leqslant \eta < b$, entonces $\int_{\eta}^{b} \varphi(x)dx < \varepsilon$. En tal caso, en virtud de la condición 1 del teorema

$$\left|\int_{\eta}^{b} f(x,y)dx\right| \leqslant \int_{\eta}^{b} |f(x,y)|dx \leqslant \int_{\eta}^{b} \varphi(x)dx < \varepsilon, \quad \eta_{\varepsilon} \leqslant \eta < b, \quad y \in Y,$$

lo que significa precisamente la convergencia uniforme de la integral $\int_{a}^{b} f(x,y)dx$ sobre el conjunto Y. □

Con ayuda del criterio de Weierstrass se establece, por ejemplo, que la integral $\int_{-\infty}^{+\infty} \frac{dx}{1+x^2+y^2}$ converge uniformemente sobre todo el eje real $-\infty < y < +\infty$.

En efecto, la integral $\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \pi$ converge, y para cualesquiera x e y se verifica la desigualdad $\frac{1}{1+x^2+y^2} \leqslant \frac{1}{1+x^2}$.

Del criterio de Cauchy para la convergencia uniforme de las funciones según un parámetro (véase el p. 39.4) se obtienen directamente las condiciones necesarias y suficientes (que también se llaman criterio de Gauchy) para la convergencia uniforme de las integrales.

Teorema 2. (criterio de Cauchy de la convergencia uniforme de las integrales). *Para que la integral* (54.1) *converja uniformemente sobre el conjunto* Y, *es necesario y suficiente que para cualquier* $\varepsilon > 0$ *exista tal* $\eta < b$, *que cualesquiera que sean* η' *y* η'', *que satisfagan las condiciones* $\eta < \eta' < b$, $\eta < \eta'' < b$, *y todo* $y \in Y$, *se verifique la desigualdad*

$$\left| \int_{\eta'}^{\eta''} f(x, y)dx \right| < \varepsilon. \tag{54.6}$$

Efectivamente, como ya se ha observado, la convergencia uniforme de la integral (54.1) es equivalente a que la función $\Phi(y, \eta)$ tienda uniformemente al límite (véase (54.5)), mientras que la desigualdad (54.6) puede escribirse, en las designaciones de (54.5), en la forma

$$|\Phi(y, \eta'') - \Phi(y, \eta')| < \varepsilon.$$

Por eso el teorema 2 no es otra cosa que simplemente una paráfrasis del teorema 4 del punto 39.4 para el caso que ahora se considera.

Ejercicios. Investíguense la convergencia y la convergencia uniforme de las integrales para todos los valores del parámetro α, indíquense las regiones de variación del parámetro α, en las que tiene lugar la convergencia uniforme de las integrales:

1. $\int_{-\infty}^{+\infty} \frac{dx}{x^2+\alpha^2}$.
2. $\int_{-\infty}^{+\infty} \frac{dx}{1+(x-\alpha)^2}$.
3. $\int_{-\infty}^{+\infty} e^{-(x-\alpha)^2} dx$.
4. $\int_0^1 \frac{dx}{x+\alpha\sqrt{x}}$.
5. $\int_0^1 \frac{dx}{x^\alpha}$.

6. Investíguese la convergencia uniforme de la integral $\int_0^{-\infty} \frac{dx}{(x-a)^p + b^p}$ cuando $a \in R$, $b \geqslant b_0 > 0$ (cuando $b > 0$, respectivamente), $p > 1$.

54.2*. CRITERIO DE LA CONVERGENCIA UNIFORME DE LAS INTEGRALES

En este punto se demostrará el criterio para la convergencia uniforme de las integrales, análogo al criterio correspondiente para la convergencia uniforme de las series (véase el p. 36.3).

Teorema 3. *Supongamos que las funciones $f(x, y)$ y $g(x, y)$ están definidas para $a \leqslant x < +\infty$, $e\ y \in Y$ (a es finito, Y es un conjunto numérico), con la particularidad de que la función $f(x, y)$ es continua respecto de la variable x, mientras que $g(x, y)$ tiene la derivada $\frac{\partial g}{\partial x}$ continua respecto de x. Si*

1) *la función $g(x, y)$ es monótona respecto de x para cada $y \in Y$, y en el conjunto Y tiende uniformemente hacia cero cuando $x \to \infty$;*

2) *la integral $\int_a^{\eta} f(x, y)dx$ está acotada como una función de las variables $\eta \in [a, +\infty)$ e $y \in Y$ en el conjunto $[a, +\infty) \times Y$;*

entonces la integral

$$\int_a^{+\infty} g(x, y)f(x, y)dx \tag{54.7}$$

converge uniformemente sobre el conjunto Y.

DEMOSTRACIÓN. De acuerdo con el segundo teorema del valor medio para las integrales (véase el p. 30.3*), para cualesquiera η' y η'', $a < \eta' < \eta''$, se verifica la igualdad

$$\int_{\eta'}^{\eta''} g(x, y)f(x, y)dx = g(\eta', y) \int_{\eta'}^{\xi} f(x, y)dx + g(\eta'', y) \int_{\xi}^{\eta''} f(x, y)dx, \tag{54.8}$$

donde $\eta' < \xi < \eta''$. En vista de la condición 2) del teorema, existe tal constante $M > 0$ que para cualesquiera $(\eta, y) \in [a, +\infty) \times Y$ tiene lugar la desigualdad $\left|\int_a^{\eta} f(x, y)dx\right| \leqslant M$. Por eso

$$\left|\int_{\eta'}^{\xi} f(x, y)dx\right| = \left|\int_a^{\xi} f(x, y)dx + \int_a^{\eta'} f(x, y)dx\right| \leqslant$$

$$\leqslant \left|\int_a^{\xi} f(x, y)dx\right| + \left|\int_a^{\eta'} f(x, y)dx\right| = 2M; \tag{54.9}$$

análogamente,

$$\left|\int_{\xi}^{\eta''} f(x,y)dx\right| \leqslant 2M. \tag{54.10}$$

Fijemos un $\varepsilon > 0$ arbitrario. En virtud de que la función $g(x,y)$ tiende uniformemente a cero en el conjunto Y, cuando $x \to +\infty$, existe tal $\eta_\varepsilon > a$ que, cualesquiera que sean $x > \eta_\varepsilon$ e $y \in Y$, se verifica la desigualdad

$$|g(x,y)| < \frac{\varepsilon}{4M}. \tag{54.11}$$

Con ayuda de las desigualdades (54.9), (54.10) y (54.11), de (54.8) se deduce que para cualesquiera $\eta' > \eta_\varepsilon$ y $\eta'' > \eta_\varepsilon$ tiene lugar la estimación

$$\left|\int_{\eta'}^{\eta''} g(x,y)f(x,y)dx\right| \leqslant$$

$$\leqslant |g(\eta',y)| \left|\int_{\eta'}^{\xi} f(x,y)dx\right| + |g(\eta'',y)| \left|\int_{\xi}^{\eta''} f(x,y)dx\right| <$$

$$< 2M\frac{\varepsilon}{4M} + 2M\frac{\varepsilon}{4M} = \varepsilon.$$

De este modo, la condición de Cauchy (véase el p. 54.1) referente a la convergencia uniforme de la integral (54.7) queda cumplida. □

OBSERVACIÓN. Se podría estimar la integral en el primer miembro de la igualdad sin recurrir al segundo teorema del valor medio, sino procediendo de manera análoga a la que se ha usado en la demostración del criterio de Dirichlet en el p. 33.6, integrarla por partes. Esto, no obstante, haría, más engorrosa la demostración sin que se eviten, en esencia, los razonamientos repetidos, propios para la demostración del segundo teorema del valor medio.

El hecho de que la función $g(x,y)$ tiene derivada continua respecto de x no es esencial y se debe sólo a que el segundo teorema del valor medio en el p. 28.3° se ha demostrado bajo las condiciones de tal suposición.

Ejemplo. La integral $\int_1^{+\infty} \frac{x \operatorname{sen} xy}{1+x^2}dx$ converge uniformemente para $y \geqslant y_0 > 0$. En efecto, la función $g(x) \overset{\text{def}}{=} \frac{x}{1+x^2}$ va decreciendo para $x \geqslant 1$ y $\lim_{x\to+\infty} g(x) = 0$, con la particularidad de que, por cuanto $g(x)$ no depende de y, dicha función tiende a cero uniformemente (cuando $x \to +\infty$) respecto de y; además

$$\left|\int_0^{\eta} \operatorname{sen} xy\,dx\right| = \frac{1-\cos\eta y}{y} \leqslant \frac{2}{y_0}.$$

De este modo, ambas condiciones del teorema 3 están cumplidas.

Problema 32. Demuéstrese que si las funciones $f(x, y)$ y $g(x, y)$ están definidas para $-\infty < a \leqslant x < +\infty$ e $y \in Y$, siendo la integral $\int_a^{+\infty} f(x, y)dx$ uniformemente convergente sobre Y, mientras que la función $g(x, y)$ es monótona respecto de x y está acotada en el conjunto $[a, +\infty) \times Y$, entonces la integral $\int_a^{+\infty} g(x, y)f(x, y)dx$ converge uniformemente sobre Y.

Ejercicios. 7. Supongamos que las funciones $f(x)$ y $g(x, y)$ son continuas respecto de x; además, cuando $x \to +\infty$, la función $g(x, y)$ tiende, de manera monótona y uniforme respecto de $y \in Y$, hacia cero y tiene derivada continua $\frac{\partial g(x, y)}{\partial x}$, $x \geqslant a$, $y \in Y$, mientras que la integral $\int_a^{+\infty} f(x)dx$ es convergente. En este caso la integral $\int_a^{+\infty} f(x)g(x, y)dx$ es uniformemente convergente sobre el conjunto Y.

Investíguese la convergencia uniforme de las integrales:

8. $\int_0^{\infty} e^{-\alpha x} \frac{\operatorname{sen}^p x}{x^q} dx$ para $\alpha \geqslant 0, p \geqslant 0, q \geqslant 0$.

9. $\int_1^{\infty} \frac{\ln^p x}{x^q} dx$ para $p \geqslant 0, q \geqslant 0$.

54.3. PROPIEDADES DE LAS INTEGRALES IMPROPIAS DEPENDIENTES DE UN PARÁMETRO

Al estudiar las propiedades de las integrales impropias dependientes de un parámetro, nos encontraremos a menudo con la permutación de los pasos límites relativos a diferentes variables. Por eso demostraremos, ante todo, un lema concerniente al problema en consideración.

Lema 1. *Sean X e Y dos conjuntos de números; la función $f(x, y)$ viene definida en su producto $X \times Y$* (véase el p. 41.2): *$x \in X$, $y \in Y$, x_0 e y_0 son unos números o infinitos cualesquiera ∞, $+\infty$, $-\infty$, y existen los límites*

$$\varphi(x) = \lim_{y \to y_0} f(x, y), \quad x \in X, \quad y \quad \psi(y) = \lim_{x \to x_0} f(x, y), \quad y \in Y.$$

Si la función f tiende uniformemente por lo menos hacia uno de los límites citados, existen y son iguales entre sí ambos límites reiterados:

$$\lim_{x \to x_0} \lim_{y \to y_0} f(x, y) = \lim_{y \to y_0} \lim_{x \to x_0} f(x, y).$$

DEMOSTRACIÓN. Supongamos, por ejemplo, que la función $f(x, y)$ tiende uniformemente en X hacia $\varphi(x)$ cuando $y \to y_0$. Entonces, para cualquier $\varepsilon > 0$ fijo existe

un entorno $U(y_0)$ tal que, cualesquiera que sean $y \in \mathring{U}(y_0) \cap Y$ *) y $x \in X$, se verifica la desigualdad

$$|f(x, y) - \varphi(x)| < \frac{\varepsilon}{2}. \tag{54.12}$$

Si $y_1 \in \mathring{U}(y_0) \cap Y$ e $y_2 \in \mathring{U}(y_0) \cap Y$, entonces

$$|f(x, y_1) - f(x, y_2)| \leqslant |f(x, y_1) - \varphi(x)| + |\varphi(x) - f(x, y_2)| < \varepsilon.$$

Pasando aquí al límite para $x \to x_0$, obtendremos

$$|\psi(y_1) - \psi(y_2)| \leqslant \varepsilon. \tag{54.13}$$

Conforme al criterio de Cauchy para la existencia del límite de una función (véase el p. 4.11), de (54.13) se deduce que existe un límite finito

$$\lim_{y \to y_0} \psi(y) = A.$$

Se ha demostrado pues la existencia del límite reiterado

$$\lim_{y \to y_0} \lim_{x \to x_0} f(x, y) = A.$$

Fijemos ahora $y_1 \in \mathring{U}(y_0) \cap Y$. En este caso de (54.12), para $y = y_1$, y de (54.13), para $y_2 \to y_0$, obtendremos, respectivamente

$$|f(x, y_1) - \varphi(x)| < \frac{\varepsilon}{2}, \quad |\psi(y_1) - A| \leqslant \varepsilon. \tag{54.14}$$

Para todo $y \in Y$ existe el límite $\lim_{x \to x_0} f(x, y) = \psi(y)$. Por ello, siendo fijo $y_1 \in \mathring{U}(y_0) \cap Y$, para $\varepsilon > 0$ dado se encontrará tal entorno $U(x_0)$ que para cualquier $x \in \mathring{U}(x_0) \cap X$ tendremos

$$|f(x, y_1) - \psi(y_1)| < \varepsilon. \tag{54.15}$$

De las desigualdades (54.14) y (54.15) para todo $x \in \mathring{U}(x_0) \cap X$ se tiene

$$|\varphi(x) - A| \leqslant |\varphi(x) - f(x, y_1)| + |f(x, y_1) - \psi(y_1)| + |\psi(y_1) - A| < 3\varepsilon,$$

lo que significa precisamente la existencia del límite reiterado

$$A = \lim_{x \to x_0} \varphi(x) = \lim_{x \to x_0} \lim_{y \to y_0} f(x, y). \quad \square$$

Teorema 4. *Sea $-\infty < a < b \leqslant +\infty$ y supongamos que la función $f(x, y)$ está definida para todo $x \in [a, b)$, $y \in Y$, y en $[a, b)$ es continua respecto de x para cualquier $y \in Y$. Entonces, si con $\eta \in [a, b)$ cualquiera la función $f(x, y)$ tiende, uniformemente en el segmento $[a, \eta]$ hacia la función $\varphi(x)$ cuando $y \to y_0$ **), y la in-*

*) Mediante $\mathring{U}$ se denota, como siempre, un entorno reducido.

**) Aquí y_0 es un número o bien uno de los infinitos ∞, $+\infty$, $-\infty$.

tegral

$$\int_a^b f(x, y)dx \tag{54.16}$$

es uniformemente convergente sobre el conjunto Y, *entonces*

$$\lim_{y \to y_0} \int_a^b f(x, y)dx = \int_a^b \lim_{y \to y_0} f(x, y)dx = \int_a^b \varphi(x)dx. \tag{54.17}$$

DEMOSTRACIÓN. Si $a < \eta < b$, de acuerdo con el teorema 2 del p. 53.1, tenemos

$$\lim_{y \to y_0} \int_a^\eta f(x, y)dx = \int_a^\eta \lim_{y \to y_0} f(x, y)dx = \int_a^\eta \varphi(x)dx. \tag{54.18}$$

Por eso, según la definición de la integral impropia, la igualdad (54.17) se puede escribir en otra forma:

$$\lim_{y \to y_0} \lim_{\eta \to b-0} \int_a^\eta f(x, y)dx = \lim_{\eta \to b-0} \lim_{y \to y_0} \int_a^\eta f(x, y)dx. \tag{54.19}$$

De este modo, resta demostrar la posibilidad de permutar el orden de los pasos límites respecto a las variables y y η para la función

$$\Phi(y, \eta) = \int_a^\eta f(x, y)dx.$$

Esto proviene del lema demostrado más arriba. En efecto, de acuerdo con (54.18), existe el límite $\lim_{y \to y_0} \Phi(y, \eta)$. Por otra parte existe también el límite

$$\lim_{\eta \to b-0} \Phi(y, \eta) = \lim_{\eta \to b-0} \int_a^\eta f(x, y)dx = \int_a^b f(x, y)dx,$$

con la particularidad de que, por hipótesis del teorema, la función tiende a su límite uniformemente en el conjunto Y. Por consiguiente, la validez de la igualdad (54.19) se desprende directamente de la afirmación del lema. □

Teorema 5. *Supongamos que la función* $f(x, y)$ *está definida y es continua (como función de dos variables) en un "rectángulo" semiabierto*

$$\{(x, y) : a \leqslant x < b, c \leqslant y \leqslant d\}, \quad -\infty < a < b \leqslant +\infty$$

$$-\infty < c < d < +\infty.$$

En este caso, si la integral $\Phi(y) = \int_a^b f(x, y)dx$ *converge uniformemente en* $[c, d]$, *será una función continua en dicho segmento.*

DEMOSTRACIÓN. Cualquiera que sea $y_0 \in [c, d]$, la función $f(x, y)$ tiende uniformemente, para $y \to y_0$, hacia la función $f(x, y_0)$ en todo segmento $[a, \eta]$, $a < \eta < b$ (véase el p. 39.4). Por ello, de acuerdo con el teorema antecedente (véase (54.17)),

$$\lim_{y \to y_0} \Phi(y) = \int_a^b \lim_{y \to y_0} f(x, y)dx = \int_a^b f(x, y_0)dx = \Phi(y_0). \quad \square$$

Teorema 6. *Cumplidas las suposiciones del teorema 5, se tiene*

$$\int_c^d \Phi(y)dy = \int_c^d dy \int_a^b f(x, y)dx = \int_a^b dx \int_c^d f(x, y)dy. \tag{54.20}$$

DEMOSTRACIÓN. Si $a < \eta < b$, entonces, según el teorema 3 del p. 53.1, tenemos

$$\int_c^d dy \int_a^\eta f(x, y)dx = \int_a^\eta dx \int_c^d f(x, y)dy. \tag{54.21}$$

La función $\Phi(y, \eta) = \int_a^\eta f(x, y)dx$ es continua respecto de y, y cuando $\eta \to b - 0$, tiende a su límite $\Phi(y)$ uniformemente en el segmento $[c, d]$. Por eso, de conformidad con el teorema 2, el p. 53.1, en el primer miembro de la igualdad (54.21) podemos pasar al límite bajo el signo de la integral para $\eta \to b - 0$:

$$\lim_{\eta \to b-0} \int_c^d dy \int_a^\eta f(x, y)dx = \lim_{\eta \to b-0} \int_c^d \Phi(y, \eta)dy =$$

$$= \int_c^d \lim_{\eta \to b-0} \Phi(y, \eta)dy = \int_c^d \Phi(y)dy = \int_c^d dy \int_a^b f(x, y)dx;$$

y en este caso el límite obtenido es finito. Por consiguiente, cuando $\eta \to b - 0$, el segundo miembro de la igualdad (54.21) tiene también el mismo límite el cual, por definición de la integral impropia, es igual a

$$\int_a^b dx \int_c^d f(x, y)dy. \quad \square$$

Demostremos un teorema sobre la permutación del orden de integración para el caso en que ambas integrales son impropias.

Teorema 7. *Supongamos que la función $f(x, y)$ está definida y es continua en el rectángulo semiabierto*

$$[(x, y) : a \leqslant x < b, c \leqslant y < d],$$

$$-\infty < a < b \leqslant +\infty, \quad -\infty < c < d \leqslant +\infty.$$

Si la integral

$$\int_a^b f(x, y)dx \tag{54.22}$$

es uniformemente convergente en cualquier segmento $[c, \eta]$, $c < \eta < d$, y la integral

$$\int_c^d f(x, y)dy \tag{54.23}$$

es uniformemente convergente en cualquier segmento $[a, \xi]$, $a < \xi < b$, *y existe, además, una de las dos integrales reiteradas*

$$\int_c^d dy \int_a^b |f(x, y)|\,dx, \quad \int_a^b dx \int_c^d |f(x, y)|\,dy,$$

existen y son iguales entre sí ambas integrales reiteradas

$$\int_c^d dy \int_a^b f(x, y)\,dx \quad y \quad \int_a^b dx \int_c^d f(x, y)\,dy,$$

es decir,

$$\int_c^d dy \int_a^b f(x, y)\,dx = \int_a^b dx \int_c^d f(x, y)\,dy. \tag{54.24}$$

DEMOSTRACIÓN. Supongamos, por ejemplo, que existe una integral

$$\int_a^b dx \int_c^d |f(x, y)|\,dy \tag{54.25}$$

y sea $c < \eta < d$. En virtud de que la integral (54.22) es uniformemente convergente en el segmento $[c, \eta]$, de acuerdo con el teorema 6, tenemos

$$\int_c^{\eta} dy \int_a^b f(x, y)\,dx = \int_a^b dx \int_c^{\eta} f(x, y)\,dy. \tag{54.26}$$

El límite del primer miembro de esta igualdad para $\eta \to d - 0$ es, obviamente, igual a

$$\int_c^d dy \int_a^b f(x, y)\,dx.$$

Probemos que el límite del segundo miembro de la igualdad (54.26) es igual a

$$\int_a^b dx \int_c^d f(x, y)\,dy,$$

es decir, en este caso resulta posible el paso límite, para $\eta \to d - 0$, bajo el signo de integral. Comprobemos si se cumplen las premisas del teorema 4 de este punto. La función $\Phi(x, \eta) = \int_c^{\eta} f(x, y)\,dy$ es continua respecto de x (véase el teorema 1 del p. 53.1) y, por hipótesis del teorema, en cualquier segmento $[a, \xi]$, $a < \xi < b$, tiende uniformemente a la integral (54.23) cuando $\eta \to d - 0$, es decir, hacia la función $F(x) = \int_c^d f(x, y)\,dy$. Por fin, la integral

$$\int_a^b \Phi(x, \eta)\,dx = \int_a^b dx \int_c^{\eta} f(x, y)\,dy$$

converge uniformemente respecto de η, $c < \eta < d$, pues

$$|\Phi(x, \eta)| \leqslant \int_c^d |f(x, y)|\,dy,$$

mientras que la integral (54.25) converge por suposición.

Por consiguiente, las condiciones del teorema 4 para el segundo miembro de la igualdad (54.26) quedan cumplidas, por lo cual

$$\lim_{\eta \to d-0} \int_a^b \Phi(x, \eta)\,dx = \int_a^b \lim_{\eta \to d-0} \Phi(x, \eta)\,dx = \int_a^b dx \int_c^d f(x, y)\,dy.$$

Así pues, la igualdad (54.24), que se demuestra, se obtiene de (54.26) pasando al límite para $\eta \to d - 0$. □

Pasemos ahora a considerar la derivabilidad de las integrales impropias dependientes de un parámetro.

Teorema 8. *Supongamos que las funciones* $f(x, y)$ *y* $\dfrac{\partial f(x, y)}{\partial y}$ *están definidas y son continuas en el rectángulo semiabierto*

$$\Delta = \{a \leqslant x < b,\ c \leqslant y \leqslant d\},$$

$$-\infty < a < b \leqslant +\infty, \quad -\infty < c < d < +\infty.$$

Si la integral $\int_a^b f(x, y)\,dx$ *converge y la integral* $\int_a^b \dfrac{\partial f(x, y)}{\partial y}\,dx$ *converge uniformemente en el segmento* $[c, d]$, *entonces la función* $\Phi(y) = \int_a^b f(x, y)\,dx$ *es continuamente derivable en dicho segmento y*

$$\frac{d}{dy} \int_a^b f(x, y)\,dx = \int_a^b \frac{\partial f(x, y)}{\partial y}\,dx.$$

DEMOSTRACIÓN. Representemos la función $\Phi(y) = \int_a^b f(x, y)\,dx$ en forma de una serie convergente en el segmento $[c, d]$:

$$\Phi(y) = \int_a^b f(x, y)\,dx = \sum_{n=1}^{\infty} \int_{\eta_n}^{\eta_{n+1}} f(x, y)\,dx, \tag{54.27}$$

donde η_n, $n = 1, 2, \ldots$, es una sucesión fija tal que

$\eta_n \in [a, b)$, $\eta_1 = a$ y $\lim\limits_{n \to \infty} \eta_n = b$, y la función $\int_a^b \dfrac{\partial f(x, y)}{\partial y}\,dx$ se representará

en forma en una serie convergente en el segmento $[c, d]$

$$\int_a^b \frac{\partial f(x, y)}{\partial y}\,dx = \sum_{n=1}^{\infty} \int_{\eta_n}^{\eta_{n+1}} \frac{\partial f(x, y)}{\partial y}\,dx. \tag{54.28}$$

De acuerdo con el teorema 4 del p. 53.2, todo término de la serie (54.28) es una derivada, respecto de la variable y, del término correspondiente de la serie (54.27), por lo cual, en virtud del teorema sobre la derivación de las series (véase el p. 36.4), la suma de la serie (54.28) es la derivada de la suma de la serie (54.27). □

Como ya se ha observado, todas las formulaciones y demostraciones anteriores se refieren a las integrales impropias dependientes de un parámetro que satisfacen las condiciones 1) y 2), enunciadas al principio del p. 54.1. De manera sumamente análoga se consideran también los casos más generales, por ejemplo, si

1') $-\infty \leqslant a < b \leqslant +\infty$.

2') para todo $y \in Y$ la función $f(x, y)$ respecto de la variable x es integrable según Riemann en cualquier segmento $[\xi, \eta]$, donde $a < \xi < \eta < b$.

La teoría construida de las integrales dependientes de un parámetro se extiende de modo natural al caso en que una integral depende de dos o, en general, de cierto número finito de parámetros $y_1, \ldots, y_n$. Con ello, muchas de las formulaciones de las definiciones y los teoremas, al igual que las demostraciones, quedan invariables, siempre que se da a las designaciones usadas el sentido nuevo. Esto atañe, por ejemplo, a la definición de la convergencia uniforme y al teorema del paso al límite bajo el signo de integral, sólo conviene considerar que $y = (y_1, \ldots, y_n)$, y_0 es un punto del mismo espacio o el infinito, mientras que $y \to y_0$ se entiende en el sentido de límite en dicho espacio.

54.4. APLICACIÓN DE LA TEORIA DE INTEGRALES DEPENDIENTES DE UN PARÁMETRO AL CÁLCULO DE LAS INTEGRALES DEFINIDAS

Hasta ahora hemos conocido dos métodos para calcular integrales definidas. El primero de ellos parte de la definición de la integral como un límite de las sumas integrales y es de amplio uso en los procedimientos numéricos de los cálculos. Se estudiará más detalladamente en el p. 60.4. El segundo método, que ya ha sido empleado constantemente, se basa sobre la búsqueda de la función subintegral primitiva y la aplicación de la fórmula de Newton — Leibniz. Resulta que a veces se logra obtener los valores exactos de las integrales definidas, utilizando la teoría de integrales dependientes de un parámetro. La ventaja de este método consiste en que en algunos casos se calculan, con ayuda de este método, las integrales de unas funciones cuyas primitivas no son funciones elementales, debido a lo cual el procedimiento usual en el que se emplea la fórmula de Newton — Leibniz resulta no aplicable.

Ejemplo 1. Supongamos que se pide calcular la integral

$$J = \int_0^1 \frac{\operatorname{arctg} x}{x\sqrt{1-x^2}}\, dx. \tag{54.29}$$

Demos a conocer los métodos de su cálculo basados sobre la sustitución de la integral dada por alguna otra que dependa de un parámetro y para la cual (54.29) sirva de valor particular.

Consideremos una función $f(x, y) = \dfrac{\operatorname{arctg} xy}{x\sqrt{1-x^2}}$ y la integral

$$J(y) = \int_0^1 f(x, y)dx = \int_0^1 \frac{\operatorname{arctg} xy}{x\sqrt{1-x^2}}\,dx. \tag{54.30}$$

Es evidente que la integral (54.29) se obtiene de aquí cuando $y = 1$. Puesto que $\frac{\operatorname{arctg} xy}{x\sqrt{1-x^2}} = 0\,(1)$ para $x \to 0$ y $\frac{\operatorname{arctg} xy}{x\sqrt{1-x^2}} = 0\left(\frac{1}{\sqrt{1-x^2}}\right)$ para $x \to 1$ y cualquier y fijo, entonces la integral (54.30) converge con todo y.

De la desigualdad $\left|\frac{\partial f(x,y)}{\partial y}\right| = \left|\frac{1}{(1+x^2y^2)\sqrt{1-x^2}}\right| \leqslant \frac{1}{\sqrt{1-x^2}}$ y de la convergencia de la integral $\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}$ se deduce que la integral

$$\int_0^1 \frac{\partial f(x,y)}{\partial y}\,dx \tag{54.31}$$

converge uniformemente sobre todo el eje real y, de conformidad con el teorema 8 del p. 54.3, es igual a $J'(y)$.

Al realizar sucesivamente los cambios de la variable de integración $x = \cos\varphi$ y $t = \operatorname{tg}\varphi$, obtendremos

$$J'(y) = \int_0^1 \frac{dx}{(1+x^2y^2)\sqrt{1-x^2}} = \int_0^{\pi/2} \frac{d\varphi}{1+y^2\cos^2\varphi} =$$

$$= \int_0^{\infty} \frac{dt}{1+y^2+t^2} = \frac{1}{\sqrt{1+y^2}} \operatorname{arctg} \frac{t}{\sqrt{1+y^2}}\Bigg|_0^{+\infty} = \frac{\pi}{2\sqrt{1+y^2}}.$$

De aquí, por definición de la integral indefinida, se infiere que

$$J(y) = \int J'(y)dy = \frac{\pi}{2} \int \frac{dy}{\sqrt{1+y^2}} = \frac{\pi}{2} \ln(y + \sqrt{1+y^2}) + C.$$

Pero, de (54.30) se deduce que $J(0) = 0$, por lo cual $C = 0$ y

$$J(y) = \frac{\pi}{2} \ln(y + \sqrt{1+y^2}).$$

Sustituyendo aquí $y = 1$, obtenemos el valor de la integral buscada (54.29)

$$J = J(1) = \frac{\pi}{2} \ln(1 + \sqrt{2}).$$

La integral (54.29) puede calcularse también aplicando la integración respecto de un parámetro. Al notar que $\frac{\operatorname{arctg} x}{x} = \int_0^1 \frac{dy}{1+x^2y^2}$, obtendremos para J la expre-

sión

$$J = \int_0^1 \frac{dx}{\sqrt{1-x^2}} \int_0^1 \frac{dy}{1+x^2y^2}. \tag{54.32}$$

La integral $\int_0^1 \frac{dx}{(1+x^2y^2)\sqrt{1-x^2}}$ converge uniformemente respecto de y, pues $\frac{1}{(1+x^2y^2)\sqrt{1-x^2}} \leqslant \frac{1}{\sqrt{1-x^2}}$, mientras que la integral $\int_0^1 \frac{dx}{\sqrt{1-x^2}}$ converge.

Por esta razón podemos cambiar en (54.32) el orden de integración (véase el teorema 6 del p. 54.3). Entonces (haciendo uso del valor, obtenido directamente más arriba, de la integral respecto de x), encontramos

$$J = \int_0^1 dy \int_0^1 \frac{dx}{(1+x^2y^2)\sqrt{1-x^2}} = \frac{\pi}{2} \int_0^1 \frac{dy}{\sqrt{1-y^2}} = \frac{\pi}{2} \ln(1+\sqrt{2}).$$

Ejemplo 2. Calculemos el valor de la integral

$$I(\alpha) = \int_0^{+\infty} \frac{\operatorname{sen} \alpha x}{x} dx. \tag{54.33}$$

Se puede mostrar que la correspondiente integral indefinida no se expresa, para $\alpha \neq 0$, en términos de las funciones elementales, a consecuencia de lo cual la integral dada no se calcula por el procedimiento habitual con ayuda de la fórmula de Newton — Leibniz.

La integral (54.33) es convergente para todo valor de α. En efecto, si $\alpha = 0$, entonces, evidentemente, $I(0) = 0$. Si, en cambio, $\alpha \neq 0$, entonces, al realizar el cambio de la variable $t = \alpha x$, se obtendrá

$$I(\alpha) = \begin{cases} \displaystyle\int_0^{+\infty} \frac{\operatorname{sen} t}{t} dt = I(1), & \text{si} \quad \alpha > 0 \\ \displaystyle\int_0^{+\infty} \frac{\operatorname{sen} t}{t} dt = -I(1), & \text{si} \quad \alpha < 0. \end{cases}$$

Por cuanto la integral $I(1)$ es convergente (véase el p. 33.5), convergerá también la integral $I(\alpha)$.

Con el fin de calcular la integral (54.33), examinemos una integral más general

$$I(\alpha, \beta) = \int_0^{+\infty} e^{-\beta x} \frac{\operatorname{sen} \alpha x}{x} dx.$$

Al derivar formalmente respecto de α bajo el signo de integral, obtendremos una integral $\int_0^{+\infty} e^{-\beta x}\cos\alpha x\,dx$, la cual converge uniformemente respecto del parámetro α, $-\infty < \alpha < +\infty$, cualquiera que sea $\beta > 0$ fijo. Por consiguiente, para $\beta > 0$ (véase el tomo 1, p. 26.4) se tiene

$$\frac{\partial I(\alpha,\beta)}{\partial\alpha} = \int_0^{+\infty} e^{-\beta x}\cos\alpha x\,dx = \frac{\beta}{\alpha^2+\beta^2},$$

de donde

$$I(\alpha,\beta) = \int_0^{\alpha}\frac{\beta\,dt}{t^2+\beta^2} + C(\beta) = \operatorname{arctg}\frac{\alpha}{\beta} + C(\beta).$$

Pero $I(0,\beta) = 0$, por lo tanto, $C(\beta) = 0$. Así pues,

$$I(\alpha,\beta) = \operatorname{arctg}\frac{\alpha}{\beta}, \quad \beta > 0.$$

Nos interesa, sin embargo, el valor de la integral $I(\alpha,\beta)$ para $\beta = 0$. Resulta más simple tratar de argumentar la posibilidad de pasar al límite bajo el signo de la integral $I(\alpha,\beta)$ cuando $\beta \to +0$. Fijemos un número $b \geqslant 0$ y mostremos que la integral $I(\alpha,\beta)$ converge uniformemente respecto del parámetro β en el segmento $[0, b]$, cualquiera que sea $\alpha \neq 0$ fijo. Efectivamente, integrando por partes (véase el p. 26.4 del tomo 1), obtendremos

$$\int_\eta^{+\infty} e^{-\beta x}\frac{\operatorname{sen}\alpha x}{x}\,dx = \frac{1}{x}e^{-\beta x}\frac{\alpha\cos\alpha x + \beta\operatorname{sen}\alpha x}{\alpha^2+\beta^2}\Bigg|_\eta^{+\infty} + \int_\eta^{+\infty} e^{-\beta x}\frac{\alpha\cos\alpha x + \beta\operatorname{sen}\alpha x}{\alpha^2+\beta^2}\frac{dx}{x^2}.$$

Elijamos η_ε de modo tal que para $\eta \geqslant \eta_\varepsilon$ se cumplan las desigualdades

$$\left|\frac{1}{\eta}e^{-\eta\beta}\frac{\alpha\cos\alpha\eta + \beta\operatorname{sen}\alpha\eta}{\alpha^2+\beta^2}\right| \leqslant \frac{|\alpha|+b\,1}{\alpha^2}\frac{}{\eta} < \frac{\varepsilon}{2},$$

$$\left|\int_\eta^{+\infty} e^{-\beta x}\frac{\alpha\cos\alpha x + \beta\operatorname{sen}x}{\alpha^2+\beta^2}\frac{dx}{x^2}\right| \leqslant \frac{|\alpha|+b}{\alpha^2}\int_\eta^{+\infty}\frac{dx}{x^2} < \frac{\varepsilon}{2}.$$

En tal caso para $\eta \geqslant \eta_\varepsilon$ obtendremos $\left|\int_\eta^{+\infty} e^{-\beta x}\frac{\operatorname{sen}\alpha x}{x}\,dx\right| < \varepsilon$, lo que precisamente demuestra la convergencia uniforme de la integral $I(\alpha,\beta)$ respecto del parámetro β en cualquier segmento $[0, b]$. Ahora, en virtud del teorema 4 del p. 54.3,

se tiene

$$I(\alpha) = I(\alpha, 0) = \lim_{\beta \to +0} I(\alpha, \beta) = \lim_{\beta \to +0} \operatorname{arctg} \frac{\alpha}{\beta} = \frac{\pi}{2} \operatorname{sign} \alpha;$$

así pues,

$$I(\alpha) = \int_0^{+\infty} \frac{\operatorname{sen} \alpha x}{x} dx = \begin{cases} \pi/2, & \text{si } \alpha > 0, \\ 0, & \text{si } \alpha = 0, \\ -\pi/2, & \text{si } \alpha < 0. \end{cases}$$

Cabe fijar la atención en que la derivación respecto de α en (54.33) llevaría a una integral divergente $\int_0^{\infty} \cos \alpha x \, dx$. La derivación se hizo posible, en $I(\alpha, \beta)$, gracias a la presencia del factor $e^{-\beta x}$, $\beta > 0$, llamado "factor de convergencia". Se denomina "método de introducción del factor de convergencia" el procedimiento por cuyo intermedio se calcula la integral de la forma $\int_0^{\infty} f(x)dx$, pasando a la integral $\int_0^{\infty} e^{-\beta x} f(x)dx$, derivando respecto de β, buscando la integral obtenida y pasando al límite cuando $\beta - 0$.

Conociendo el valor de $I(\alpha)$ es posible hallar con facilidad también los valores de varias integrales semejantes. Por ejemplo, podemos mostrar fácilmente (de lo que haremos uso en adelante) que

$$\int_{-\infty}^{+\infty} \frac{1 - \cos \alpha x}{x^2} dx = |\alpha| \pi. \tag{54.34}$$

Efectivamente, integrando por partes, encontraremos

$$\int_{-\infty}^{+\infty} \frac{1 - \cos \alpha x}{x^2} dx = \alpha \int_{-\infty}^{+\infty} \frac{\operatorname{sen} \alpha x}{x} dx = 2\alpha \int_0^{+\infty} \frac{\operatorname{sen} \alpha x}{x} dx = |\alpha| \pi.$$

Ejercicios. Calcúlense las integrales

10. $\displaystyle\int_0^a \frac{\ln(1 + ax)}{1 + x^2} dx \; (a > 0).$

11. $\displaystyle\int_0^{\pi/2} \frac{\ln(1 + a \cos x)}{\cos x} dx \; (|a| < 1).$

12. $\displaystyle\int_0^{\infty} \frac{\operatorname{arctg} ax}{x(1 + b^2 x^2)} dx.$

13. $\int_0^{\infty} \frac{\cos ax + \cos bx - 2}{x^2}\,dx \ (a, b \neq 0).$

14. $\int_0^{\infty} e^{-ax}\frac{\operatorname{sen}^2 bx}{x^2}\,dx.$

54.5. INTEGRALES DE EULER

Examinemos las integrales

$$B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx, \tag{54.35}$$

$$\Gamma(s) = \int_0^{+\infty} x^{s-1}e^{-x}dx, \tag{54.36}$$

llamadas *integrales de Euler de primera y segunda especie*, respectivamente, la integral (54.35) lleva el nombre de *función beta* y la (54.36), *función gamma*.

Aclaremos, ante todo, para qué valores de los parámetros p, q y s tienen sentido los segundos miembros de las fórmulas (54.35) y (54.36). Consideremos al principio la integral (54.35). En el caso general la función subintegral tiene dos peculiaridades: cuando $x = 0$ y cuando $x = 1$, por lo cual representemos la integral en la forma

$$B(p, q) = \int_0^{1/2} x^{p-1}(1-x)^{q-1}dx + \int_{1/2}^1 x^{p-1}(1-x)^{q-1}dx.$$

Al comparar la primera integral en el segundo miembro con la integral $\int_0^{1/2} x^{p-1}dx$, y la segunda integral, con $\int_{1/2}^1 (1-x)^{q-1}dx$, que son convergentes para $p > 0$ y $q > 0$, respectivamente, y divergentes para $p \leqslant 0$ y $q \leqslant 0$, respectivamente, (véase el p. 33.3), llegamos a que el dominio de definición de la función beta (54.35) en el plano p, q será el ángulo recto $p > 0$, $q > 0$.

Luego, la integral $B(p, q)$ converge uniformemente en cada ángulo recto $p \geqslant p_0, q \geqslant q_0$, cualesquiera que sean $p_0 > 0$ y $q_0 > 0$. En efecto, de acuerdo con el criterio de Weierstrass, esto se desprende (véase el p. 54.1) de la desigualdad

$$x^{p-1}(1-x)^{q-1} \leqslant x^{p_0-1}(1-x)^{q_0-1}, \quad 0 \leqslant x \leqslant 1,$$

y de la convergencia, demostrada más arriba, de la integral

$$B(p_0, q_0) = \int_0^1 x^{p_0-1}(1-x)^{q_0-1}dx, \quad p_0 > 0, \quad q_0 > 0.$$

Por cuanto todo punto (p, q), $p > 0$, $q > 0$, pertenece a cierto ángulo $p > p_0$, $q > q_0$, siendo elegidos de modo adecuado los números $p_0 > 0$ y $q_0 > 0$, entonces, en virtud del teorema 5 del p. 54.3, la función $B(p, q)$ es continua en todo su dominio de definición.

Con el fin de encontrar el dominio de definición de la función gamma (54.36), representémosla en la forma

$$\Gamma(s) = \int_0^1 x^{s-1}e^{-x}dx + \int_1^{+\infty} x^{s-1}e^{-x}dx. \qquad (54.37)$$

Al comparar el primer sumando en el segundo miembro con la integral $\int_0^1 x^{s-1}dx$, la cual converge para $s > 0$ y diverge para $s \leqslant 0$, llegamos a que la integral $\int_0^1 x^{s-1} \times e^{-x}dx$ converge y diverge para los mismos valores del parámetro s. En lo que se refiere a la segunda integral en el miembro derecho de la igualdad (54.37), converge para cualquier valor de s. Esto proviene, por ejemplo, de la validez, para todo s, de la igualdad $x^{s-1}e^{-x} = o(e^{-x/2})$ cuando $x \to +\infty$, y de la convergencia de la integral $\int_1^{+\infty} e^{-x/2}dx = 2e^{-1/2}$. De este modo, la integral (54.36) converge para todo $s > 0$ y diverge cuando $s \leqslant 0$.

Probemos ahora que la integral (54.36) converge uniformemente en todo segmento $[s_1, s_2]$, donde $0 < s_1 < s_2 < +\infty$. En efecto, sea $s_1 \leqslant s \leqslant s_2$; entonces, si $0 \leqslant x \leqslant 1$, se tiene

$$x^{s-1}e^{-x} \leqslant x^{s_1-1}e^{-x},$$

y si $x \geqslant 1$, se tiene

$$x^{s-1}e^{-x} \leqslant x^{s_2-1}e^{-x}$$

y como las integrales $\int_0^1 x^{s_1-1}e^{-x}dx$ y $\int_1^{+\infty} x^{s_2-1}e^{-x}dx$ convergen, de la fórmula (54.37) se deduce, en virtud del criterio de Weierstrass para la convergencia uniforme de integrales (véase el p. 54.1), la convergencia uniforme de la integral $\Gamma(s)$ en el segmento $[s_1, s_2]$. De aquí proviene, en vista del teorema 5 del p. 54.3, que la función $\Gamma(s)$ es continua en todo su dominio de definición.

Ejercicio 15. Demuéstrese que las funciones $B(p, q)$ y $\Gamma(s)$ son infinitamente derivables.

Problema 33. Demuéstrese que $B(p, q)$ y $\Gamma(s)$ son las funciones analíticas.

Demos a conocer algunas propiedades de las integrales $\Gamma(s)$ y $B(p, q)$. Ante todo, de la fórmula (54.36) se obtiene directamente

$$\Gamma(s) > 0 (s > 0), \qquad (54.38)$$

en particular, la función gamma no tiene ceros. Luego, al integrar por partes, obtenemos

$$\Gamma(s+1) = \int_0^{+\infty} x^s e^{-x}dx = -x^s e^{-x}\Big|_0^{+\infty} + s\int_0^{+\infty} x^{s-1}e^{-x}dx = s\Gamma(s). \qquad (54.39)$$

De este modo, si $s > n$ ($n = 1, 2, \ldots$), entonces

$$\Gamma(s) = (s-1)(s-2)\ldots(s-n)\Gamma(s-n). \tag{54.40}$$

Para cualquier $s > 0$ puede elegirse un número entero no negativo n de modo tal que sea $0 < s - n \leqslant 1$ ($n = 0, 1, 2, \ldots$), y en este caso $\Gamma(s)$ se expresará, mediante la fórmula (54.40), en términos del valor de la función gamma en cierto punto del intervalo (0, 1]. En otras palabras, si se conoce el valor de la función gamma en el intervalo (0,1], se puede hallar su valor en cualquier punto.

Hemos de notar, además, que $\Gamma(1) = \int_0^{+\infty} e^{-x}dx = 1$, y, por lo tanto, en virtud de la fórmula (54.40),

$$\Gamma(n+1) = n!.$$

De aquí se ve que la función gamma $\Gamma(s+1)$ es una prolongación de la función $s!$, definida sólo para $s = 0, 1, 2, \ldots$ enteros, a todo el semieje $s > -1$ de números reales.

Demostraremos las siguientes propiedades de la función beta $B(p, q)$.

1. Para cualesquiera $p > 0$ y $q > 0$

$$B(p, q) = B(q, p). \tag{54.41}$$

Para convencerse de esto, basta efectuar el cambio de la variable $t = 1 - x$ en la integral (54.35).

2. Para cualesquiera $p > 0$ y $q > 1$

$$B(p, q) = \frac{q-1}{p+q-1} B(p, q-1). \tag{54.42}$$

Análogamente, en virtud de la simetría (véase (54.41)), para cualesquiera $q > 0$ y $p > 1$

$$B(p, q) = \frac{p-1}{p+q-1} B(p-1, q). \tag{54.43}$$

En efecto, al integrar por partes (54.35) y observar que $x^p(1-x)^{q-2} = x^{p-1}(1-x)^{q-2} - x^{p-1}(1-x)^{q-1}$, obtendremos $B(p, q) =$

$$= \int_0^1 (1-x)^{q-1} d\frac{x^p}{p} = \frac{x^p(1-x)^{q-1}}{p}\Big|_0^1 + \frac{q-1}{p}\int_0^1 x^p(1-x)^{q-2}dx =$$

$$= \frac{q-1}{p}\int_0^1 x^{p-1}(1-x)^{q-2}dx - \frac{q-1}{p}\int_0^1 x^{p-1}(1-x)^{q-1}dx =$$

$$= \frac{q-1}{p} B(p, q-1) - \frac{q-1}{p} B(p, q),$$

de donde se desprende (54.42), y, en virtud de la simetría, también (54.43).

3. Para cualesquiera $p > 0$

$$B(p, n) = B(n, p) = \frac{1 \cdot 2 \cdot 3 \ldots (n-1)}{p(p+1)\ldots(p+n-1)}, \quad n = 1, 2, \ldots .$$

Esta fórmula se obtiene como resultado de la aplicación sucesiva de la correlación (54.42), si se toma en consideración que $B(p, 1) = \int_0^1 x^{p-1}dx = \frac{1}{p}$. Si $p = m$ es también un número natural, entonces $B(m, n) = \frac{(n-1)!(m-1)!}{(m+n-1)!}$.

Entre las funciones $B(p, q)$ y $\Gamma(s)$ existe una correlación que se establece mediante la fórmula de Euler

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p > 0, \quad q > 0. \tag{54.44}$$

Demostrémosla, siguiendo el método de Dirichlet. Realicemos en la fórmula (54.36) el cambio de la variable $x = (1+t)y$, $t > 0$:

$$\frac{\Gamma(s)}{(1+t)^s} = \int_0^{+\infty} y^{s-1}e^{-(1+t)y}dy$$

y pongamos $s = p + q$, $p > 0$, $q > 0$; entonces

$$\frac{\Gamma(p+q)}{(1+t)^{p+q}} = \int_0^{+\infty} y^{p+q-1}e^{-(1+t)y}dy.$$

Multipliquemos ambos miembros de esta igualdad por t^{p-1} e integremos respecto de t entre 0 y $+\infty$:

$$\Gamma(p+q)\int_0^{+\infty} \frac{t^{p-1}}{(1+t)^{p+q}}dt = \int_0^{+\infty} t^{p-1}dt \int_0^{+\infty} y^{p+q-1}e^{-(1+t)y}dy. \tag{54.45}$$

Realicemos el cambio de variable $t = \frac{x}{1-x}$ en la integral que figura en el primer miembro de la igualdad (54.45):

$$\int_0^{+\infty} \frac{t^{p-1}}{(1+t)^{p+q}}dt = \int_0^1 x^{p-1}(1-x)^{q-1}dx = B(p, q). \tag{54.46}$$

Para calcular el segundo miembro de la igualdad observemos que

$$\int_0^{+\infty} t^{p-1}dt \int_0^{+\infty} y^{p+q-1}e^{-(1+t)y}dy =$$

$$= \lim_{\xi \to +0} \int_0^{+\infty} t^{p-1} dt \int_\xi^{+\infty} y^{p+q-1} e^{-(1+t)y} dy. \quad (54.47)$$

En efecto, designando $\Phi(t, \xi) = \int_\xi^{+\infty} y^{p+q-1} e^{-(1+t)y} dy$, de la estimación

$$0 \leqslant \Phi(t, 0) - \Phi(t, \xi) \leqslant \int_0^\xi y^{p+q-1} e^{-y} dy,$$

$$\int_0^{+\infty} t^{p-1} \Phi(t, \xi) dt \leqslant \int_0^{+\infty} t^{p-1} \Phi(t, 0) dt$$

concluimos que para $\xi \to +0$ la función $\Phi(t, \xi)$ tiende hacia $\Phi(t, 0)$ uniformemente respecto de $t \in (0, +\infty)$ y que la integral $\int_0^{+\infty} t^{p-1} \Phi(t, \xi) dt$ converge uniformemente respecto de ξ, pues converge la integral (54.45). Por consiguiente, en el segundo miembro de (54.47) podemos pasar al límite bajo el signo de la integral exterior.

Luego,

$$\int_0^{+\infty} t^{p-1} dt \int_\xi^{+\infty} y^{p+q-1} e^{-(1+t)y} dy =$$

$$= \int_\xi^{+\infty} y^{p+q-1} e^{-y} dy \int_0^{+\infty} t^{p-1} e^{-ty} dt, \quad \xi > 0, \quad p \geqslant 1, \quad q \geqslant 1. \quad (54.48)$$

La permutación del orden de integración es posible aquí, porque, primero, la integral $t^{p-1} \int_\xi^{+\infty} y^{p+q-1} e^{-(1+t)y} dy$ es uniformemente convergente respecto de t en cualquier segmento $[0, a]$, lo que se deduce de la estimación uniforme de la función subintegral

$$t^{p-1} y^{p+q-1} e^{-(1+t)y} \leqslant a^{p-1} y^{p+q-1} e^{-y}, \quad 0 \leqslant t \leqslant a,$$

y de la convergencia de la integral $\int_\xi^{+\infty} y^{p+q-1} e^{-y} dy$, segundo, la integral

$$y^{p+q-1} e^{-y} \int_0^\infty t^{p-1} e^{-ty} dt$$

converge uniformemente respecto de y en cualquier segmento $[\xi, b]$, $\xi > 0$, lo que se deduce de la estimación uniforme de la función subintegral

$$y^{p+q-1} e^{-y} t^{p-1} e^{-ty} \leqslant b^{p+q-1} t^{p-1} e^{-\xi t}$$

y de la convergencia de la integral $\int_0^{+\infty} t^{p-1}e^{-\xi t}dt$; tercero, la integral que figura en el segundo miembro de la igualdad (54.48) existe. De este modo, la legitimidad del cambio del orden de integración en (54.48) se infiere del teorema 7, p. 54.3 (observemos que en este caso la función subintegral es no negativa).

Al realizar el cambio de variable $ty = u$, obtendremos

$$\int_\xi^{+\infty} y^{p+q-1}e^{-y}dy \int_0^{+\infty} t^{p-1}e^{-t}dt = \Gamma(p) \int_\xi^{+\infty} y^{q-1}e^{-y}dy. \qquad (54.49)$$

Por fin,

$$\lim_{\xi \to +0} \int_\xi^{+\infty} y^{q-1}e^{-y}dy = \Gamma(q). \qquad (54.50)$$

De (54.45)—(54.50) obtenemos la fórmula (54.44) para $p \geqslant 1$, $q \geqslant 1$. Si ahora $p > 0$ y $q > 0$, entonces, según lo demostrado, se tiene

$$B(p + 1, q + 1) = \frac{\Gamma(p + 1)\Gamma(q + 1)}{\Gamma(p + q + 2)}.$$

Aplicando las correlaciones (54.39), (54.42) y (54.43), obtendremos la fórmula (54.44) bajo el supuesto de que $p > 0$, $q > 0$. □

54.6. FUNCIONES DE VALORES COMPLEJOS DE UN ARGUMENTO REAL

En lo que sigue se considerarán sistemáticamente las funciones de valores complejos $w(t) = u(t) + iv(t)$ del argumento real t (las funciones $u(t)$ y $v(t)$ adquieren valores reales). Ya nos encontramos con los conceptos de límite y continuidad de las funciones similares. La derivada de la función $w(t)$ se determina según la fórmula

$$w'(t) \stackrel{\text{def}}{=} u'(t) + iv'(t).$$

Mostremos, por ejemplo, que, de acuerdo con dicha regla, $(e^{i\alpha t})' = i\alpha e^{i\alpha t}$. En efecto,

$$(e^{i\alpha t})' = (\cos \alpha t + i \operatorname{sen} \alpha t)' = -\alpha \operatorname{sen} \alpha t + i\alpha \cos \alpha t =$$
$$= i\alpha(\cos \alpha t + i \operatorname{sen} \alpha t) = i\alpha e^{i\alpha t}$$

Análogamente se determina también la integral (propia o impropia) de la función $w = u + iv$:

$$\int_a^b w(t)dt = \int_a^b u(t)dt + i\int_a^b v(t)dt, \quad -\infty \leqslant a < b \leqslant +\infty.$$

La integral $\int_a^b (u(x) + iv(x))dx$ se denomina impropia, si lo es por lo menos una

de las integrales $\int_a^b u(x)dx$ y $\int_a^b v(x)dx$. Además, la integral impropia $\int_a^b (u(x) + iv(x))dx$ se llama *convergente*, si convergen tanto $\int_a^b u(x)dx$, como también $\int_a^b v(x)dx$.

En este caso

$$\int_a^b (u(x) + iv(x))dt \stackrel{\text{def}}{=} \int_a^b u(x)dx + i\int_a^b v(x)dx.$$

La función w se denomina *absolutamente integrable*, si lo son las funciones u y v.

Es evidente que toda una serie de las propiedades de las integrales de las funciones reales (linealidad de la integral, aditividad de ésta según los conjuntos, etc.) se extienden automáticamente a las funciones de valores complejos. Observemos, por ejemplo, que si $w(x) = u(x) + iv(x)$, donde $u(x)$ y $v(t)$ son funciones reales integrables según Riemann en el segmento $[a, b]$, entonces la integral $\int_a^b w(x)dx$ es también el límite de las sumas integrales $\sigma_\tau = \sum_{i=1}^{k} w(\xi_i)\Delta x_i (\tau = \{x_i\}_{i=0}^{i=k}$ es la partición del segmento $[a, b], x_{i-1} \leqslant \xi_i \leqslant x_i, \Delta x_i = x_i - x_{i-1}, i = 1, 2, \ldots, k$). De aquí, al igual que para las funciones reales, se desprende que en este caso la función $|w(x)|$ es también integrable según Riemann y que se verifica la desigualdad

$$\left|\int_a^b w(x)dx\right| \leqslant \int_a^b |w(x)|dx.$$

Esta desigualdad es válida también para las funciones de valores complejos absolutamente integrables en el sentido impropio, lo que se establece por medio de un paso límite.

Mas, en el caso de las funciones que toman valores complejos se debe tener cuidado al recurrir a los análogos de los teoremas demostrados para las funciones reales. La razón para ello es que no todas las afirmaciones válidas para las funciones del argumento real, que toman sólo valores reales, se extienden a las funciones de valores complejos. Con una situación semejante ya nos hemos encontrado al estudiar las funciones vectoriales (véase el p. 15.2 y también 37.9*). Por ejemplo, las afirmaciones, semejantes al teorema de Rolle y, por lo tanto, al de Lagrange de valores medios no son válidas para las funciones de valores complejos. Esto lo demuestra el ejemplo, aducido en el p. 15.2, si se escribe en términos de los valores complejos.

A saber, consideremos una función $f(t) = \cos t + i \operatorname{sen} t$, $0 \leqslant t \leqslant 2\pi$; para ella se verifica $f(0) = f(2\pi) = 1, f'(t) = -\operatorname{sen} t + i \cos t$. Por cuanto $|f'(t)| = \sqrt{\operatorname{sen}^2 t + \cos^2 t} = 1$, no existe tal punto $\xi \in [0, 2\pi]$, que se verifique $f'(\xi) = 0$. Por consiguiente, el análogo del teorema de Rolle no tiene lugar en el caso dado.

Tampoco resulta cierta la regla de L'Hospital cuya demostración se ha basado sobre el teorema del valor medio. Probemos esto con un ejemplo *).

Sea $f(t) = t$, $g(t) = t + t^2 e^{\frac{i}{t^2}}$, $0 < t < 1$. Por cuanto, de acuerdo con la fórmula de Euler, $e^{i/t^2} = \cos\frac{1}{t^2} + i \operatorname{sen}\frac{1}{t^2}$, entonces

$$|e^{i/t^2}| = \sqrt{\cos^2\frac{1}{t^2} + \operatorname{sen}^2\frac{1}{t^2}} = 1.$$

Por eso, $\lim\limits_{t\to 0} f(t) = \lim\limits_{t\to 0} g(t) = 0$, y

$$\lim_{t\to 0} \frac{f(t)}{g(t)} = \lim_{t\to 0} (1 + te^{i/t^2}) = 1. \tag{54.51}$$

Teniendo presente que

$$g'(t) = 1 + \left(2t - \frac{2i}{t}\right)e^{i/t^2}, \quad 0 < t < 1,$$

obtenemos

$$|g'(t)| \geqslant \left|\frac{2i}{t} - 2t\right| - 1 \geqslant \frac{2}{t} - 2t - 1 \geqslant \frac{2}{t} - 1 = \frac{2-t}{t}.$$

Por consiguiente, $\left|\dfrac{f'(t)}{g'(t)}\right| = \dfrac{1}{|g'(t)|} \leqslant \dfrac{t}{2-t}$, a consecuencia de lo cual

$$\lim_{t\to 0} \frac{f'(t)}{g'(t)} = 0. \tag{54.52}$$

Comparando (54.51) y (54.52), nos convencemos de que en el caso dado la regla de L'Hospital no es aplicable.

54.7*. COMPORTAMIENTO ASINTÓTICO DE LA FUNCIÓN GAMMA

Mostremos que el comportamiento asintótico de la función gamma

$$\Gamma(s+1) = \int_0^{+\infty} e^{-x} x^s dx, \quad s > -1, \tag{54.53}$$

puede ser descrito, para valores suficientemente grandes de la variable independiente s, mediante una fórmula bastante sencilla que contiene sólo funciones elementales.

*) Este ejemplo se ha tomado del libro de W. Rudin "Fundamentos del análisis matemático". (W. Rudin, *Principles of mathematical analysis*, 2nd ed., New York-Toronto-London, 1964).

Es fácil ver que la función subintegral en la integral (54.53) toma su valor máximo cuando $x = s$. Realicemos en esta integral el cambio de la variable de integración, trasladando el punto $x = s$ al nuevo origen de coordenadas: $x = s + y$, y realizando, a continuación, la transformación de semejanza con un coeficiente igual a s: $y = st$, es decir, pongamos $x = s(1 + t)$. Obtendremos

$$\Gamma(s + 1) = e^{-s}s^{s+1} \int_{-0}^{+\infty} [e^{-t(1+t)}]^s dt. \tag{54.54}$$

Examinemos la función

$$\varphi(t) \stackrel{\text{def}}{=} e^{-t}(1 + t), \quad -\infty < t < +\infty. \tag{54.55}$$

Por cuanto $\varphi'(t) = -te^{-t}$, la función φ decrece cuando $t < 0$ y crece cuando $t < 0$; en el punto $t = 0$ la función alcanza su valor máximo $\varphi(0) = 1$. Luego, al poner

$$h(t) \stackrel{\text{def}}{=} -t + \ln(1 + t), \quad -1 < t < -\infty \tag{54.56}$$

obtendremos

$$\varphi(t) = e^{h(t)}, \quad -1 < t < +\infty, \tag{54.57}$$

donde para $|t| < 1$

$$h(t) = -\frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \ldots$$

y por eso

$$h(t) = -\frac{t^2}{2} + o(t^2), \quad t \to 0. \tag{54.58}$$

Así pues, la función gamma puede ser representada en la forma (véanse (54.54), (54.55) y (54.57))

$$\Gamma(s + 1) = e^{-s}s^{s+1} \int_{-1}^{+\infty} e^{sh(t)} dt, \tag{54.59}$$

donde el comportamiento de la función $h(t)$ para $t \to 0$ se describe por la correlación (54.58).

Antes de pasar a la deducción de la fórmula asintótica para $\Gamma(s + 1)$ con $s \to +\infty$, aclaremos el método de su obtención con ayuda de unos razonamientos que no son muy rigurosos, pero verosímiles. La gráfica de la función $\varphi(t)$ es de la forma expuesta en la fig. 229. A medida que crece el parámetro s, la gráfica de la función $[\varphi(t)]^s$ irá "apretándose" contra el eje de la variable t y el segmento unidad del eje de ordenadas. Está claro por eso que la integral

$$\int_{-1}^{+\infty} e^{sh(t)} dt, \tag{54.60}$$

en el segundo miembro de la fórmula (54.59) se aproximará perfectamente bien para

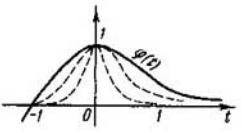


Fig. 229

grandes valores de s, mediante la integral

$$\int_{-\delta}^{\delta} e^{sh(t)}dt, \tag{54.61}$$

donde $\delta > 0$ es arbitrario, pero fijo, con la particularidad de que, con una exactitud tanto mayor cuanto mayor sea el valor del parámetro s. En otras palabras, si s es suficientemente grande, entonces tanto para $-1 < t < -\delta$, como para $t > \delta$, los valores de la función $e^{sh(t)}$ son tan pequeños que cada una de las integrales $\int_{-1}^{-\delta} e^{sh(t)}dt$ y $\int_{\delta}^{\infty} e^{sh(t)}dt$ puede ser despreciada con alto grado de precisión. Es natural esperar que, siendo fijado $\delta > 0$, el error relativo que se obtiene como resultado de aproximar la integral (54.60) con ayuda de las integrales del tipo (54.61), también puede hacerse tan pequeño como se quiera, a cuenta de la elección del parámetro s lo suficientemente grande.

En virtud de (54.58), al tomar $\delta > 0$ suficientemente pequeño, podemos aproximar con éxito la integral (54.61) por medio de la integral

$$\int_{-\delta}^{\delta} e^{-\frac{st^2}{2}}dt = \sqrt{\frac{2}{s}} \int_{-\delta\sqrt{\frac{s}{2}}}^{\delta\sqrt{\frac{s}{2}}} e^{-u^2}du. \tag{54.62}$$

Si $\delta > 0$, el segundo miembro de esta igualdad tiende, para $s \to +\infty$, hacia la integral de Poisson (véase el p. 48.2)

$$\int_{-\infty}^{+\infty} e^{-u^2}du = \sqrt{\pi}. \tag{54.63}$$

De resultas, para los valores grandes de s la integral (54.60) resulta ser, en cierto sentido, bien aproximada por la expresión $\sqrt{2\pi/s}$ (véanse (54.62) y (54.63)). Por esta razón es natural tratar de demostrar una igualdad asintótica

$$\int_{-1}^{+\infty} e^{sh(t)}dt \sim \sqrt{\frac{2\pi}{s}}, \; s \to +\infty.$$

Mostremos que esta igualdad realmente tiene lugar. Elijamos arbitrariamente ε, $0 < \varepsilon < \frac{1}{2}$. En virtud de (54.58), existe tal δ, $0 < \delta < 1$, que para cualesquiera $t \in [-\delta, \delta]$ se cumple la desigualdad

$$\left| h(t) + \frac{t^2}{2} \right| < \varepsilon t^2,$$

es decir,

$$-\left(\frac{1}{2} + \varepsilon\right) t^2 < h(t) < -\left(\frac{1}{2} - \varepsilon\right) t^2.$$

Consecuentemente (por ser monótona la función e^x), para cualesquiera $s > 0$ se verifica la desigualdad

$$e^{-\frac{(1+2\varepsilon)st^2}{2}} < e^{sh(t)} < e^{-\frac{(1-2\varepsilon)st^2}{2}}.$$

Integrándola en el segmento $[-\delta, \delta]$, obtendremos

$$\int_{-\delta}^{\delta} e^{-\frac{(1+2\varepsilon)st^2}{2}} dt \leqslant \int_{-\delta}^{\delta} e^{sh(t)} dt \leqslant \int_{-\delta}^{\delta} e^{-\frac{(1-2\varepsilon)st^2}{2}} dt. \quad (54.64)$$

Estimemos ahora en cuánto la integral (54.61), que figura en el medio de esta desigualdad, se diferencia de la integral (54.60) que nos interesa. Recordando que la función $\varphi(t) = e^{sh(t)} = e^{-t}(1 + t)$ (véanse (54.55) y (54.57)) crece en el intervalo $[-1, -\delta]$ y decrece en $[\delta, +\infty)$, obtenemos para todos los $s > 1$:

$$0 < \int_{-1}^{+\infty} e^{sh(t)} dt - \int_{-\delta}^{\delta} e^{sh(t)} dt =$$

$$= \int_{-1}^{-\delta} e^{(s-1)h(t)} e^{h(t)} dt + \int_{\delta}^{+\infty} e^{(s-1)h(t)} e^{h(t)} dt \leqslant$$

$$\leqslant e^{(s-1)h(-\delta)} \int_{-1}^{-\delta} e^{h(t)} dt + e^{(s-1)h(\delta)} \int_{\delta}^{+\infty} e^{h(t)} dt \leqslant$$

$$\leqslant [e^{(s-1)h(-\delta)} + e^{(s-1)h(\delta)}] \int_{-1}^{+\infty} e^{h(t)} dt \leqslant C_1 e^{-\alpha_1 s}, \quad (54.65)$$

donde

$$\alpha_1 = -\max\{h(-\delta), h(\delta)\} > 0,$$

$$C_1 = (e^{-h(-\delta)} + e^{-h(\delta)}) \int_{-1}^{+\infty} e^{h(t)} dt < +\infty.$$

Observemos que la función $h(t)$ (véase (54.56)) alcanza un máximo estricto en el punto $t = 0$, con la particularidad de que $h(0) = 0$; por ello, $h(-\delta) < 0$ y $h(\delta) < 0$.

De la manera semejante se estiman también las integrales extremas en la desigualdad (54.64). Al realizar el cambio de la variable de integración

$u = t\sqrt{\frac{(1+2\varepsilon)s}{2}}$, obtendremos (véase (54.63))

$$\int_{-\infty}^{+\infty} e^{-\frac{(1+2\varepsilon)st^2}{2}}\,dt = \sqrt{\frac{2}{(1+2\varepsilon)s}}\int_{-\infty}^{+\infty} e^{-u^2}du = \sqrt{\frac{2\pi}{(1+2\varepsilon)s}}.$$

Ahora, por analogía con (54.65), tendremos

$$0 < \sqrt{\frac{2\pi}{(1+2\varepsilon)s}} - \int_{-\delta}^{\delta} e^{-\frac{(1+2\varepsilon)st^2}{2}}\,dt =$$

$$= \int_{-\infty}^{+\infty} e^{-\frac{(1+2\varepsilon)st^2}{2}}\,dt - \int_{-\delta}^{\delta} e^{-\frac{(1+2\varepsilon)st^2}{2}}\,dt =$$

$$= \int_{-\infty}^{-\delta} e^{-\frac{(1+2\varepsilon)(s-1)t^2}{2}} e^{-\frac{(1+2\varepsilon)t^2}{2}}\,dt +$$

$$+ \int_{\delta}^{+\infty} e^{-\frac{(1+2\varepsilon)(s-1)t^2}{2}} e^{-\frac{(1+2\varepsilon)t^2}{2}}\,dt \leqslant$$

$$\leqslant e^{-\frac{(1+2\varepsilon)(s-1)\delta^2}{2}}\left(\int_{-\infty}^{-\delta} e^{-\frac{(1+2\varepsilon)t^2}{2}}\,dt + \int_{\delta}^{+\infty} e^{-\frac{(1+2\varepsilon)t^2}{2}}\,dt\right) =$$

$$= e^{-\frac{(1+2\varepsilon)(s-1)\delta^2}{2}}\sqrt{\frac{2}{1+2\varepsilon}}\left(\int_{-\infty}^{-\delta\sqrt{\frac{1+2\varepsilon}{2}}} e^{-u^2}du + \int_{\delta\sqrt{\frac{1+2\varepsilon}{2}}}^{+\infty} e^{-u^2}du\right) \leqslant$$

$$\leqslant e^{-\frac{(1+2\varepsilon)(s-1)\delta^2}{2}}\sqrt{\frac{2}{1+2\varepsilon}}\int_{-\infty}^{+\infty} e^{-u^2}\,du \leqslant$$

$$\leqslant e^{-\frac{(1+2\varepsilon)(s-1)\delta^2}{2}}\sqrt{2\pi} \leqslant C_2 e^{-\alpha_2 s}, \quad (54.66)$$

donde $\alpha_2 = \dfrac{(1+2\varepsilon)\delta^2}{2} > 0$, $C_2 = e^{\frac{(1+2\varepsilon)\delta^2}{2}}\sqrt{2\pi}$.

Del mismo modo se obtiene también la estimación

$$0 < \sqrt{\frac{2}{(1-2\varepsilon)s}} - \int_{-\delta}^{\delta} e^{-\frac{(1-2\varepsilon)st^2}{2}}dt \leqslant C_3 e^{-\alpha_3 s}, \qquad (54.67)$$

donde $\alpha_3 = \dfrac{(1-2\varepsilon)\delta^2}{2} > 0$, $C_3 = e^{\frac{(1-2\varepsilon)\delta^2}{2}}\sqrt{2\pi}$.

Al poner $\alpha = \min\{\alpha_1, \alpha_2, \alpha_3\}$ y sustituir (54.65), (54.66) y (54.67) en (54.64), obtendremos, para las constantes correspondientes $C_4 > 0$ y $C_5 > 0$ (dependientes de ε):

$$\sqrt{\frac{2\pi}{(1+2\varepsilon)s}} + C_4 e^{-\alpha s} \leqslant \int_{-1}^{+\infty} e^{sh(t)}dt \leqslant \sqrt{\frac{2\pi}{(1-2\varepsilon)s}} + C_5 e^{-\alpha s}.$$

Dividamos la desigualdad obtenida por $\sqrt{2\pi/s}$

$$\frac{1}{\sqrt{1+2\varepsilon}} + C_4 e^{-\alpha s}\sqrt{\frac{s}{2\pi}} \leqslant$$

$$\leqslant \frac{1}{\sqrt{\frac{2\pi}{s}}} \int_{-1}^{+\infty} e^{sh(t)}dt \leqslant \frac{1}{\sqrt{1-2\varepsilon}} + C_5 e^{-\alpha s}\sqrt{\frac{s}{2\pi}}.$$

Por consiguiente, pasando al límite para $s \to +\infty$, tendremos, cualquiera que sea $\varepsilon > 0$:

$$\frac{1}{\sqrt{1+2\varepsilon}} \leqslant \lim_{s\to+\infty} \frac{1}{\sqrt{\frac{2\pi}{s}}} \int_{-1}^{+\infty} e^{sh(t)}dt \leqslant \frac{1}{\sqrt{1-2\varepsilon}}.$$

Al hacer tender aquí ε hacia cero, obtendremos

$$\lim_{s\to+\infty} \frac{1}{\sqrt{\frac{2\pi}{s}}} \int_{-1}^{+\infty} e^{sh(t)}dt = 1,$$

o, que es lo mismo, la igualdad asintótica buscada

$$\int_{-1}^{+\infty} [e^{-t}(1+t)]^s dt \sim \sqrt{\frac{2\pi}{s}}, \quad s \to +\infty.$$

Al multiplicar ambos miembros de la igualdad por $e^{-s}s^{s+1}$, en virtud de (54.54), obtendremos la fórmula asintótica

$$\Gamma(s+1) \sim \sqrt{2\pi}\, e^{-s} s^{s+\frac{1}{2}}, \quad s \to +\infty, \tag{54.68}$$

llamada *fórmula de Stirling para la función gamma*. Esta fórmula es, evidentemente, una generalización de la fórmula de Stirling para el factorial de números naturales (véase el p. 37.8), que se obtiene de (54.68), si ponemos en ésta $s = n$, pues $\Gamma(n+1) = n!$ (véase el p. 54.5).

54.8*. SERIES ASINTÓTICAS

En el p. 37.10* estudiábamos los desarrollos de las funciones en series de potencias asintóticas cuando $x \to +\infty$. Recordemos que la serie

$$a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \ldots + \frac{a_n}{x^n} + \ldots$$

se denomina *desarrollo asintótico de la función f* para $x \to +\infty$, si las sumas parciales de dicha serie

$$S_n(x) = a_0 + \frac{a_1}{x} + \ldots + \frac{a_n}{x^n}$$

satisfacen la condición

$$f(x) - S_n(x) = o\left(\frac{1}{x^n}\right), \quad x \to +\infty.$$

El concepto de desarrollo asintótico de una función se generaliza del modo natural a las series según los sistemas de funciones que forman las así llamadas *sucesiones asintóticas*.

Definición 3. *Una sucesión de funciones $\varphi_n(x)$, $n = 0, 1, 2, \ldots$, definidas en cierto entorno reducido del punto a (finito o infinito) lleva el nombre de sucesión asintótica para $x \to a$, si para cualquier $n = 0, 1, 2, \ldots$, tiene lugar la correlación*

$$\varphi_{n+1}(x) = o(\varphi_n(x)), \quad x \to a. \tag{54.69}$$

Son ejemplos de sucesiones asintóticas para $x \to a$ las expresiones $\varphi_n(x) = (x-a)^n$, si a es un punto finito, y $\varphi_n(x) = x^{-n}$, si $a = +\infty$ o bien $a = -\infty$, $n = 0, 1, 2, \ldots$.

Definición 4. *Sea $\varphi_n(x)$, $n = 0, 1, 2, \ldots$, una sucesión asintótica para $x \to a$. La serie*

$$a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_n\varphi_n(x) + \ldots \tag{54.70}$$

se llama serie asintótica (o desarrollo asintótico) para $x \to a$ de la función dada f, definida en cierto entorno reducido del punto a, si sus sumas parciales

$$S_n(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_n\varphi_n(x) \tag{54.71}$$

satisfacen la condición: para cualquier $n = 0, 1, 2, \ldots$ *se verifica la igualdad asintótica*

$$f(x) - S_n(x) = o(\varphi_n(x)), \quad x \to a. \tag{54.72}$$

Lema 2. *Sea* $\varphi_n(x)$, $n = 0, 1, 2, \ldots$, *una sucesión asintótica para* $x \to a$. *Para que la serie* (54.70) *sea un desarrollo asintótico de la función* f *para* $x \to a$, *es necesario y suficiente que se verifique*

$$f(x) - \overline{S_n}(x) = O(\varphi_{n+1}(x)), \quad x \to a, \quad n = 0, 1, 2, \ldots . \tag{54.73}$$

En otras palabras, la serie (54.70) es un desarrollo asintótico de la función f para $x \to a$ cuando, y sólo cuando, su suma parcial $S_n(x)$ sirva de valor aproximado de la función $f(x)$ con la exactitud de hasta $O(\varphi_{n+1}(x))$ para $x \to a$, es decir, el orden del error no supere el orden del primer término que se desecha.

DEMOSTRACIÓN DE LA NECESIDAD DE LA CONDICIÓN (54.73). La correlación (54.72) para $n = 1, 2, \ldots$ puede escribirse en la forma

$$f(x) - S_{n-1}(x) - a_n\varphi_n(x) = o(\varphi_n(x)), \quad x \to a,$$

de donde

$$f(x) - S_{n-1}(x) = a_n\varphi_n(x) + o(\varphi_n(x)) = O(\varphi_n(x)), \quad x \to a, \quad n = 1, 2, \ldots$$

es decir, se cumple la condición (54.73). □

DEMOSTRACIÓN DE LA SUFICIENCIA DE LA CONDICIÓN (54.73). En virtud de (54.73) y (54.69), tenemos

$$f(x) - S_n(x) = O(\varphi_{n+1}(x)) = O(o(\varphi_n(x))) = o(\varphi_n(x)), \quad x \to a,$$
$$n = 0, 1, 2, \ldots ,$$

lo que coincide con (54.72). □

Resulta curioso notar que si para cualquier $n = 0, 1, 2, \ldots$, se cumple la condición

$$f(x) - S_n(x) = O(\varphi_n(x)), \quad x \to a, \tag{54.74}$$

más débil comparada con la (54.72), se deduce de ella, en virtud de (54.69), la igualdad asintótica (54.72). En otras palabras, el cumplimiento de la condición (54.74) para todo $n = 0, 1, 2, \ldots$ significa que la serie (54.70) es un desarrollo asintótico de la función f para $x \to a$. En efecto, de (54.74) tenemos para $n = 1, 2, \ldots$,

$$f(x) - S_{n-1}(x) = a_n\varphi_n(x) + O(\varphi_n(x)) = O(\varphi_n(x)) =$$
$$= O(o(\varphi_{n-1}(x))) = o(\varphi_{n-1}(x)), \quad x \to a,$$

es decir, la condición (54.72).

Si la sucesión asintótica $\varphi_n(x)$, $n = 0, 1, 2, \ldots$, es de tal índole que existe un entorno reducido del punto a, en el que para todo $n = 0, 1, 2, \ldots$ tiene lugar la desigualdad $\varphi_n(x) \neq 0$, entonces por analogía con el caso de las series asintóticas de potencias de las funciones obtenemos:

si una función f *se desarrolla para* $x \to a$ *en la serie asintótica* (54.70), *este desarrollo es único y sus coeficientes se determinan sucesivamente según las fórmulas*

$$a_0 = \lim_{x \to a} f(x), \quad a_n = \lim_{x \to a} \frac{1}{\varphi_n(x)} \left[f(x) - \sum_{k=0}^{n-1} a_k \varphi_k(x) \right],$$

$$n = 1, 2, \ldots .$$

No obstante, para la búsqueda práctica de los desarrollos asintóticos de las funciones dadas dicha fórmula no es siempre cómoda. Resulta más fácil a menudo obtener el desarrollo necesario por otro medio, en particular, en el caso de las integrales, integrando por partes. En este último caso la sucesión asintótica $\{\varphi_n(x)\}$ no se predetermina, por regla general, de antemano, sino que se construye, partiendo de las propiedades de la función dada en el entorno del punto a.

Ejemplo. Desarrollemos en serie asintótica, para $x \to +\infty$, la función

$$F(x, \alpha) = \int_x^{+\infty} \frac{e^{it}}{t^\alpha} dt, \quad x > 0, \tag{54.75}$$

($\alpha > 0$ es un parámetro), eligiendo la sucesión asintótica correspondiente. Por cuanto

$$F(x, \alpha) = \int_x^{+\infty} \frac{\cos t}{t^\alpha} dt + i \int_x^{+\infty} \frac{\operatorname{sen} t}{t^\alpha} dt,$$

entonces, de acuerdo con el criterio de Dirichlet (véase el p. 33.6), las partes imaginaria y real de la función $F(x, \alpha)$ representan en sí, para $x > 0$, las integrales convergentes. Por eso converge también la integral (54.75). Ha de notarse que las partes real e imaginaria de la integral $\frac{1}{2} F\left(x^2, \frac{1}{2}\right)$ las constituyen las integrales incompletas de Fresnel (véase § 34)

$$\int_x^{+\infty} \cos \theta^2 d\theta, \quad \int_x^{+\infty} \operatorname{sen} \theta^2 d\theta.$$

Con el fin de convencerse de esto, basta realizar el cambio de la variable de integración $t = \theta^2$ en la integral $\frac{1}{2} F\left(x^2, \frac{1}{2}\right)$.

Integrando por partes (54.75), obtendremos

$$F(x, \alpha) = \int_x^{+\infty} \frac{e^{it}}{t^\alpha} dt = \frac{ie^{ix}}{x^\alpha} - i\alpha \int_x^{+\infty} \frac{e^{it}}{t^{\alpha+1}} dt = \frac{ie^{ix}}{x^\alpha} - i\alpha F(x, \alpha + 1).$$

Aplicando sucesivamente esta fórmula a los valores de la función F, que se obtienen en el segundo miembro, tendremos

$$F(x, \alpha) = \frac{ie^{ix}}{x^\alpha} - i\alpha F(x, \alpha + 1) = \frac{ie^{ix}}{x^\alpha} - i\alpha \left[\frac{ie^{ix}}{x^{\alpha+1}} - i(\alpha + 1) F(x, \alpha + 2) \right] =$$

$$= \frac{ie^{ix}}{x^{\alpha}} - \frac{\alpha i^2 e^{ix}}{x^{\alpha+1}} + \frac{\alpha(\alpha+1)i^3 e^{ix}}{x^{\alpha+2}} + \ldots + \frac{(-1)^n \alpha(\alpha+1)\ldots(\alpha+n-1)i^{n+1}e^{ix}}{x^{\alpha+n}} +$$

$$+ (-i)^{n+1}\alpha(\alpha+1)\ldots(\alpha+n)F(x, \alpha+n+1) =$$

$$= \frac{ie^{ix}}{x^{\alpha}} \sum_{k=0}^{n} \frac{\alpha(\alpha+1)\ldots(\alpha+k-1)}{(ix)^k} + \frac{\alpha(\alpha+1)\ldots(\alpha+n)}{i^{n+1}} \times$$

$$\times F(x, \alpha+n+1). \quad (54.76)$$

La serie

$$\frac{ie^{ix}}{x^{\alpha}} \sum_{n=0}^{\infty} \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{(ix)^n} \quad (54.77)$$

es un desarrollo asintótico de la función $F(x, \alpha)$ para $x \to +\infty$. En efecto, la sucesión de funciones $\varphi_n(x) = e^{ix}x^{-n-\alpha}$, $n = 0, 1, \ldots$, es, como es fácil de comprobar, asintótica, y para las sumas parciales $S_n(x, \alpha)$ de la serie (54.77) tenemos, en virtud de (54.76):

$$|F(x, \alpha) - S_n(x, \alpha)| = \left| \frac{\alpha(\alpha+1)\ldots(\alpha+n)}{i^{n+1}} F(x, \alpha+n+1) \right| =$$

$$= \alpha(\alpha+1)\ldots(\alpha+n) \left| \int_x^{+\infty} \frac{e^{it}}{t^{\alpha+n+1}} dt \right| \leqslant$$

$$\leqslant \alpha(\alpha+1)\ldots(\alpha+n) \int_x^{+\infty} \frac{dt}{t^{\alpha+n+1}} =$$

$$= \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{x^{\alpha+n}} = O\left(\frac{e^{\alpha}}{x^{\alpha+n}}\right), \quad x \to \infty,$$

es decir, se cumple la condición (54.74), y, por tanto, la serie (54.77) es, de hecho, un desarrollo asintótico de la función $F(x, \alpha)$ para $x \to +\infty$.

54.9*. DESARROLLO ASINTÓTICO DE LA FUNCIÓN GAMMA INCOMPLETA

Cualquiera que sea $x > 0$, para la función gamma $\Gamma(s)$ se tiene

$$\Gamma(s) = \int_0^{+\infty} t^{s-1}e^{-t}dt = \int_0^{x} t^{s-1}e^{-t}dt + \int_x^{+\infty} t^{s-1}e^{-t}dt.$$

La función

$$\Gamma(s,x) \stackrel{\text{def}}{=} \int_x^{+\infty} t^{s-1}e^{-t}dt, \quad x > 0, \tag{54.78}$$

se denomina función gamma incompleta. Está definida para todos los valores reales del parámetro s. Hallemos su desarrollo asintótico para $x \to +\infty$. Realizando en el segundo miembro de (54.78) la integración por partes, obtendremos

$$\Gamma(s,x) = \int_x^{+\infty} t^{s-1}e^{-t}dt =$$

$$= x^{s-1}e^{-t} + (s-1)\int_x^{+\infty} t^{s-2}e^{-t}dt = x^{s-1}e^{-t} + (s-1)\Gamma(s-1,x).$$

Aplicando sucesivamente esta fórmula a los valores de la función gamma incompleta, que se obtienen en el segundo miembro, tendremos

$$\Gamma(s,x) = x^{s-1}e^{-x} + (s-1)x^{s-2}e^{-x} + \ldots$$

$$\ldots + (s-1)(s-2)\ldots(s-n+1)x^{s-n}e^{-x} + (s-1)(s-2)\ldots$$

$$\ldots (s-n)\Gamma(s-n,x) = e^{-x}x^s \sum_{k=0}^{n} \frac{(s-1)(s-2)\ldots(s-k+1)}{x^k} +$$

$$+ (s-1)(s-2)\ldots(s-n)\Gamma(s-n,x).$$

De aquí, para $n > s-1$ tenemos

$$\left|\Gamma(s,x) - e^{-x}x^s \sum_{k=0}^{n} \frac{(s-1)(s-2)\ldots(s-k+1)}{x^k}\right| =$$

$$= |(s-1)(s-2)\ldots(s-n)\Gamma(s-n,x)| \leqslant$$

$$\leqslant |(s-1)\ldots(s-n)| \int_x^{+\infty} \frac{e^{-t}}{t^{n-s+1}}dt \leqslant$$

$$\leqslant |(s-1)(s-2)\ldots(s-n)| \frac{1}{x^{n-s-1}} \int_x^{+\infty} e^{-t}dt =$$

$$= |(s-1)\ldots(s-n)| \frac{e^{-x}}{x^{n-s-1}} = O\left(\frac{e^{-x}}{x^{n-s-1}}\right), \quad x \to +\infty,$$

es decir, para las sumas parciales de la serie

$$e^{-x}x^s \sum_{n=0}^{\infty} \frac{(s-1)(s-2)\dots(s-n+1)}{x^n} \tag{54.79}$$

y para la sucesión $\varphi_n(x) = x^{-n+s+1}e^{-x}$, que es asintótica (lo que se comprueba con facilidad), la condición (54.73) se cumple cuando $x \to +\infty$. De este modo, la serie (54.79) es un desarrollo asintótico de la función gamma incompleta $\Gamma(s, x)$ para $x \to +\infty$.

En el p. 54.7° se ha encontrado el primer término del desarrollo asintótico de la función gamma $\Gamma(s+1)$ para $s \to +\infty$. Se pueden hallar también los términos siguientes, es decir, desarrollar la función gamma en una serie asintótica. Esta serie tendrá la siguiente forma

$$\Gamma(s+1) \sim (2\pi)^{\frac{1}{2}} e^{-s} s^{s+\frac{1}{2}} \left(1 + 3c_3 \frac{2!}{112^2}\frac{2}{s} + \right.$$
$$\left. + 5c_5 \frac{4!}{2!2^4}\left(\frac{2}{s}\right)^2 + \dots\right), \quad s \to +\infty. \tag{54.80}$$

Aquí $\{c_k\}$ es una sucesión de coeficientes del desarrollo en una serie de potencias (en el entorno de cero) de la función $t = t(z)$ definida mediante la igualdad $\frac{1}{2}z^2 = -h(t)$, donde $h(t)$ viene dada por la fórmula (54.56).

Puede obtenerse también el desarrollo asintótico para el logaritmo natural de la función gamma. El desarrollo citado tiene por expresión

$$\ln\Gamma(s) \sim \left(s - \frac{1}{2}\right)\ln s - s + \frac{1}{2}\ln 2\pi + \sum_{n=1}^{\infty} \frac{B_{2n}}{2n(2n-1)s^{2n-1}},$$
$$s \to +\infty \tag{54.81}$$

y se denomina *serie de Stirling*. Aquí, B_{2n} son los así llamados *números de Bernoulli* que se definen mediante la igualdad

$$\sum_{k=0}^{m-1} k^n = \frac{1}{n+1}\sum_{j=0}^{n} C_{n+1}^{j} B_j m^{n+1-j}$$

(todos los números de Bernoulli impares, salvo $B_1 = -\frac{1}{2}$, son nulos).

De la fórmula (54.81) podemos obtener (por potenciación) el desarrollo asintótico para la función gamma en el que los coeficientes serán expresados en forma explícita. El desarrollo tiene por expresión

$$\Gamma(s) \sim (2\pi)^{\frac{1}{2}} e^{-s} s^{s-\frac{1}{2}} \left\{1 + \frac{1}{12s} + \frac{1}{288s^2} - \frac{139}{51840s^3} + \dots\right\}, \quad s \to +\infty.$$

La demostración de las fórmulas (54.80) y (54.81) sale de los márgenes de este libro. La descripción de los métodos por cuyo intermedio se obtienen los desarrollos semejantes se dan en el libro de M. V. Fedoriuk "Método del punto de ensilladura" M., 1977.

54.10. OBSERVACIONES SOBRE LAS INTEGRALES MÚLTIPLES DEPENDIENTES DE UN PARÁMETRO

Hemos considerado más arriba las integrales "unidimensionales" dependientes de un parámetro, es decir, el caso en que tanto la variable de integración como el parámetro eran variables numéricas. Esta teoría se generaliza para el caso de las integrales múltiples que dependen de un parámetro "unidimensional" o "multidimensional", es decir, de las integrales de la forma

$$F(y) = \int f(x, y)dG. \tag{54.82}$$

Aquí la función $f(x, y)$ está definida respecto de la variable x en el conjunto abierto $G \subset R^n$ y es integrable según Riemann en cualquier conjunto D abierto y medible según Jordan de tal género que $\overline{D} \subset G$. El parámetro y recorre cierto conjunto Y, el cual puede representarse, por ejemplo, por un subconjunto del espacio m-dimensional R^m, mientras que la integral (54.82) se entiende, en el caso general, en el sentido impropio.

La integral (54.82) se denomina *convergente*, si para todo $y_0 \in Y$ fijo converge la integral

$$\int f(x, y_0)dG.$$

Cuando $n \geqslant 2$, esto, como se sabe (véase el p. 48.3), es equivalente a la condición de convergencia de la integral

$$\int |f(x, y_0)|\, dG.$$

A la integral convergente (54.82) (como también a toda sucesión de conjuntos abiertos medibles según Jordan G_k, $k = 1, 2, \ldots$, que agota de manera monótona el conjunto G) se le confronta del modo natural una serie de cuya suma sirve la misma:

$$\int f(x, y)dG = \int f(x, y)dG_1 + \sum_{k=1}^{\infty} \int f(x, y)d(G_{k+1} \setminus \overline{G}_k). \tag{54.83}$$

Por analogía con el caso unidimensional se define también la integral uniformemente convergente.

Definición 5. *La integral convergente* (54.82) *se llama uniformemente convergente, si para todo* $\varepsilon > 0$ *existe un compacto* $A \subset G$ *tal que para cualquier conjunto* D, *abierto y medible según Jordan, para el cual* $A \subset D \subset \overline{D} \subset G$ *se verifica la desigualdad*

$$\left| \int f(x, y)d(G \setminus \Gamma) \right| < \varepsilon.$$

Esta definición es equivalente a la siguiente:

Definición 5'. *La integral convergente* (54.82) *se llama uniformemente convergente, si, cualquiera que sea la sucesión de conjuntos abiertos y medibles según Jordan* G_k, $k = 1, 2, \ldots$, *que agota de manera monótona el conjunto abierto* G, *y el número* $\varepsilon > 0$, *existe un número* k_ε, *dependiente de la sucesión dada y del número* ε, *tal que para todo número* $k \geqslant k_\varepsilon$ *y todos los* $y \in Y$ *se verifica la desigualdad*

$$\left| \int f(x, y)d(G \setminus \overline{G}_k) \right| < \varepsilon.$$

Si la integral (54.82) converge uniformemente en el conjunto G con relación al parámetro $y \in Y$, entonces la serie (54.83) es también uniformemente convergente en G.

Para las integrales múltiples dependientes de un parámetro quedan vigentes los teoremas sobre su continuidad, derivabilidad e integrabilidad, análogos a los demostrados anteriormente. Esta afirmación se comprueba con facilidad y no será el objeto de nuestra consideración detallada.

Se encuentran integrales que dependen de un parámetro de un modo más complejo: en las integrales citadas no sólo la función subintegral f depende de un parámetro, sino también el conjunto G, según el cual se realiza la integración, es decir, $G = G(y)$:

$$F(y) = \int\limits_{G(y)} f(x, y)dx. \tag{54.84}$$

A título de ejemplo de tal integral en el caso unidimensional sirve la integral

$$F(y) = \int\limits_a^b \frac{dx}{|x - y|^\alpha}, \quad a \leqslant y \leqslant b.$$

Aquí $G(y)$ se compone de dos (salvo el caso $y = a$ e $y = b$) intervalos (a, y) y (y, b) que varían con el cambio del parámetro y.

Examinemos un ejemplo análogo en el espacio n-dimensional. Sea G un conjunto abierto en R^n, y supongamos que $\mu = \mu(x)$ es continua en G, $\rho = \rho(x, y)$ constituye la distancia entre los puntos x e y, $x \in G$, $y \in R^n$, y α es un cierto número. Las integrales de la forma

$$u(y) = \int \frac{\mu(x)dG}{\rho^\alpha} \tag{54.85}$$

se llaman *potenciales* y se relacionan al tipo (54.84), puesto que en calidad de conjunto, según el cual se realiza la integración, sirve en dichas integrales el conjunto $G \setminus \{y\}$, dependiente de y (el dominio de integración en la fórmula (54.85) se ha denotado, como lo hacemos siempre, simplemente con G). Si $\alpha = 1$ y $n = 3$, la función (54.85) recibe el nombre de *potencial newtoniano*.

Problema 34. Demuéstrese que si G es un conjunto abierto medible según Jordan y la función $\mu = \mu(x)$ es continua en su clausura $\overline{G}$, entonces la integral (54.85) será, para $\alpha < n$, continua en todo el espacio.

CAPÍTULO VII

SERIES DE FOURIER. INTEGRAL DE FOURIER

§ 55. SERIES TRIGONOMÉTRICAS DE FOURIER

55.1. DEFINICIÓN DE LA SERIE DE FOURIER. PLANTEAMIENTO DE LOS PROBLEMAS FUNDAMENTALES

Definición 1. *La serie de la forma*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx, \tag{55.1}$$

donde $\{a_n\}$, $\{b_n\}$ *son sucesiones de números reales, se denomina serie trigonométrica.*

Sus sumas parciales son combinaciones lineales de las funciones que componen el sistema

$$1) \cos x, \operatorname{sen} x, \cos 2x, \operatorname{sen} 2x, \ldots, \cos nx, \operatorname{sen} nx, \ldots . \tag{55.2}$$

Definición 2. *El conjunto de las funciones* (55.2) *se llama sistema trigonométrico.*

Lema 1. *El sistema trigonométrico* (55.2) *posee las siguientes propiedades:*

1) *Una integral, extendida al segmento* $[-\pi, \pi]$, *del producto de dos funciones distintas que integran el sistema es nula (esta propiedad lleva el nombre de ortogonalidad* *) *del sistema* (55.2)), *es decir,*

$$\begin{gathered} \int_{-\pi}^{\pi} \cos nx \cos mx\,dx = 0, \quad n \neq m, \\ \int_{-\pi}^{\pi} \operatorname{sen} nx \operatorname{sen} mx\,dx = 0, \quad n \neq m, \\ \int_{-\pi}^{\pi} \cos nx \operatorname{sen} mx\,dx = 0, \quad m, n = 0, 1, 2, \ldots; \end{gathered} \tag{55.3}$$

$$2) \int_{-\pi}^{\pi} \cos^2 nx\,dx = \int_{-\pi}^{\pi} \operatorname{sen}^2 nx\,dx = \pi, \quad n = 1, 2, \ldots . \tag{55.4}$$

---

*) El origen del término "ortogonalidad" se explicará en el p. 58.1.

DEMOSTRACIÓN. Para cualesquiera m, n enteros y no negativos, siendo $m \neq n$, tenemos

$$\int_{-\pi}^{\pi} \operatorname{sen} nx \operatorname{sen} mx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} [\cos(n-m)x - \cos(n+m)x]\,dx =$$

$$= \frac{\operatorname{sen}(n-m)x}{2(n-m)}\Bigg|_{-\pi}^{\pi} - \frac{\operatorname{sen}(n+m)x}{2(n+m)}\Bigg|_{-\pi}^{\pi} = 0.$$

De modo análogo se demuestran también las otras dos igualdades (55.3).

Demostremos ahora (55.4):

$$\int_{-\pi}^{\pi} \cos^2 nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} (1 + \cos 2nx)\,dx = \pi,$$

$$\int_{-\pi}^{\pi} \operatorname{sen}^2 nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} (1 - \cos 2nx)\,dx = \pi. \quad \square$$

Teorema 1. *Sea*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx \tag{55.5}$$

y supongamos que la serie en el segundo miembro de esta igualdad converge uniformemente en el segmento $[-\pi, \pi]$. *Entonces*

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx,$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx, \qquad n = 1, 2, \ldots. \tag{55.6}$$

DEMOSTRACIÓN. Por cuanto la serie, que figura en el segundo miembro de la igualdad (55.5), converge uniformemente en el segmento $[-\pi, \pi]$, mientras que todos sus términos representan funciones continuas en dicho segmento, entonces la suma $f(x)$ de la serie es continua en el segmento $[-\pi, \pi]$ y la propia serie puede ser integrada término a término (véase el p. 36.4) entre los límites de $-\pi$ a π:

$$\int_{-\pi}^{\pi} f(x)\,dx = \int_{-\pi}^{\pi} \left(\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx\right) dx =$$

$$= \frac{a_0}{2}\int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos nx\,dx + b_n \int_{-\pi}^{\pi} \operatorname{sen} nx\,dx = \pi a_0.$$

De aquí proviene la primera de las fórmulas (55.6).

Si la serie (55.5) se multiplica, término a término, por $\cos nx$ y $\operatorname{sen} nx$ ($n = 1, 2, \ldots$), las series obtenidas serán también uniformemente convergentes en el segmento $[-\pi, \pi]$ (véase la propiedad 2° en el p. 36.3).

Integrando término a término estas series y utilizando la propiedad de ortogonalidad (55.3) del sistema trigonométrico y las igualdades (55.4), tendremos

$$\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \int_{-\pi}^{\pi} a_n \cos^2 nx\,dx = \pi a_n,$$

$$\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx = \int_{-\pi}^{\pi} b_n \operatorname{sen}^2 nx\,dx = \pi b_n.$$

De las correlaciones obtenidas se infiere directamente la fórmula (55.6). □

Hemos de notar ahora que las integrales (55.6) tienen sentido no sólo para las funciones continuas en el segmento $[-\pi, \pi]$, sino también, por ejemplo, para funciones, la integral de las cuales converge absolutamente en este segmento.

Recordemos que el concepto de integral absolutamente convergente (como también el de integral simplemente convergente) se ha introducido sólo para las funciones definidas en cierto intervalo (a, b), $-\infty \leqslant a < b \leqslant +\infty$, que tienen tal número finito de puntos x_i, $i = 0, 1, 2, \ldots, k$, $a \leqslant x_0 < x_1 < \ldots < x_k \leqslant b$, que la función f es integrable según Riemann en cualquier segmento $[\xi_i, \eta_i]$, donde $x_{i-1} < \xi_i < \eta_i < x_i$. En este caso, si $a = -\infty$, entonces $x_0 = -\infty$, y si $b = +\infty$, entonces $x_k = +\infty$. Los números $x_0, x_1, \ldots, x_k$ se denominan *puntos singulares de la función f*.

Si, asumidas estas suposiciones, la integral $\int_a^b |f(x)|\,dx$ converge, siempre tiene sentido y converge también la integral $\int_a^b f(x)dx$ (véase el p. 33.5).

Las funciones se llaman *absolutamente integrables* en un segmento, si la integral del valor absoluto de dichas funciones converge en el mismo segmento.

Observemos que si una función es integrable según Riemann en cierto segmento, su valor absoluto es también integrable según Riemann en el mismo (véase el p. 28.1) y, por consiguiente, una función integrable según Riemann en un segmento es absolutamente integrable en él.

Si la integral de la función f converge absolutamente en el segmento $[-\pi, \pi]$, todas las integrales (55.6) para dicha función son también convergentes, puesto que representan en sí las integrales del producto de la función absolutamente integrable $f(x)$ por una función acotada (seno o coseno) y, de acuerdo con el lema 2 en el p. 33.5, tales integrales son absolutamente convergentes.

Definición 3. *Supongamos que la función $f(x)$ es absolutamente integrable en el segmento $[-\pi, \pi]$. La serie trigonométrica* (55.1), *cuyos coeficientes se determinan por la fórmula (55.6), se denomina serie de Fourier* [*]) *o, en forma más detallada, serie trigonométrica de Fourier, mientras que los números a_n y b_n se llaman coeficientes de Fourier de la función f.*

*) J. B. Fourier (1768—1830), físico y matemático francés.

En este caso se escribe

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

Las sumas parciales de orden n de esta serie se designarán mediante $S_n(x, f)$ o, de forma más breve, $S_n(x)$. Subrayemos que aquí el signo $\sim$ no es una igualdad asintótica, sino simplemente una correspondencia: a una función se le asigna su serie de Fourier.

El teorema 1 puede parafrasearse en términos mencionados del modo siguiente.

Toda serie trigonométrica uniformemente convergente es la serie de Fourier de su suma.

Ejercicios. 1. Supongamos que la función f es absolutamente integrable en el segmento $[-\pi, \pi]$ y sea

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

En este caso, si la función f es par, se tiene $b_n = 0, n = 1, 2, \ldots$, si, en cambio, f es una función impar, entonces $a_n = 0, n = 0, 1, 2, \ldots$.

2. ¿Será la serie trigonométrica

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n^2}$$

una serie de Fourier?

En este párrafo se estudiarán las funciones periódicas, es decir, tales funciones f, para cada una de las cuales existe un $T > 0$ tal que, cualquiera que sea x perteneciente al dominio de definición de la función f, los valores $x + T$ y $x - T$ pertenecen también a este campo y se verifica la igualdad

$$f(x + T) = f(x).$$

Las funciones de este género se llaman T-periódicas.

Ejercicio 3. Muéstrese que la función f que es igual a cero en todo punto racional y a uno en todos los puntos irracionales, tiene como su período cualquier número racional y ninguno de los números irracionales puede ser su período.

Sea f absolutamente integrable en el segmento $[-\pi, \pi]$ y, por consiguiente, se le puede asignar una serie de Fourier. Si la serie citada converge en cierto conjunto, converge hacia una función 2π-periódica, puesto que todos sus términos son 2π-periódicos. Resulta pues cómodo que la propia función f sea "prolongada periódicamente" con el período de 2π. Las comillas se deben a que en la realidad la función f puede ser prolongada periódicamente sólo en el caso en que $f(-\pi) = f(\pi)$.

Si esta condición no se cumple, llamaremos *prolongación de la función* f una función 2π-periódica $\bar{f}$ que se obtendrá bajo el supuesto de que para cualquier pun-

to $x \in [-\pi, \pi)$ en el que queda definida la función f (recordemos que en virtud de la integrabilidad absoluta de la función f en el segmento $[-\pi, \pi]$ ésta queda definida en todos los puntos del segmento, a excepción, quizás, de su conjunto finito); tenemos

$$\bar{f}(x + 2\pi k) = f(k), \quad k = 0, \pm 1, \pm 2, \ldots .$$

Para el caso cuando $f(-\pi) \neq f(\pi)$, tal prolongación conduce a que los valores de las funciones f y $\bar{f}$, para $x = \pi$, no coinciden. No obstante, por cuanto los coeficientes de Fourier de la función se determinan con ayuda de las integrales (55.6), esto no llevará a su cambio y, por lo tanto, las series de Fourier de la función dada f y de la prolongada $\bar{f}$ coinciden.

Diremos que en la mencionada prolongación periódica la función $\bar{f}$ puede no ser continua en los puntos $2\pi k$, $k = 0, \pm 1, \pm 2, \ldots$, si la función f es continua para $x = -\pi$ y $x = \pi$. La función prolongada f será continua en los puntos $2\pi k$, si f es continua en $x = -\pi$ y $x = \pi$, con la particularidad de que $f(-\pi) = f(\pi)$. La continuidad en otros puntos se conserva cuando tiene lugar la prolongación periódica: si f es continua en el punto $x \in (-\pi, \pi)$, $\bar{f}$ será continua en cualquier punto $x + 2k\pi$, $k = 0, \pm 1, \pm 2, \ldots$.

La función prolongada $\bar{f}$ se denotará con frecuencia mediante el mismo símbolo f que caracteriza la función prolongable.

Si la función f es 2π-periódica, al calcular sus coeficientes de Fourier (véase (55.6)), la integración puede realizarse en cualquier segmento de longitud 2π, por ejemplo, en el segmento $[0, 2\pi]$:

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x)dx,$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx\,dx, \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \operatorname{sen} nx\,dx.$$

Efectivamente, si una función φ tiene período igual a T y para cierto número $a \in R$ es integrable sobre el segmento $[a, a + T]$, entonces, cualquiera que sea $b \in R$, será integrable también en el segmento $[b, b + T]$, con la particularidad de que

$$\int_b^{b+T} \varphi(x)dx = \int_a^{a+T} \varphi(x)dx,$$

es decir, la integral $\int_a^{a+T} \varphi(x)dx$ no depende de cómo se elige el número $a \in R$. Esta propiedad de las funciones periódicas se demuestra fácilmente por cambio de variable de integración y recomendamos que el mismo lector lleve a cabo la demostración.

En el § 58 generalizaremos el concepto de serie trigonométrica de Fourier, a saber, definiremos y estudiaremos las series de Fourier referidas a un sistema ortogonal arbitrario de funciones. Ahora, en el párrafo presente, sólo se estudiarán las se-

ries trigonométricas de Fourier de las funciones absolutamente integrables (véase también el p. 58.6).

Ante todo se estudiará la cuestión sobre las condiciones que garantizan la convergencia de la serie de Fourier. En el caso de que la serie de Fourier de una función dada $f(x)$ converja bajo ciertas condiciones, se aclarará a qué es igual su suma $S(x)$ y, en particular, cuando coincide con la función $f(x)$. Se estudiará la "velocidad" de convergencia de las series de Fourier y las condiciones de que ésta depende. Se mostrará que incluso en el caso cuando la serie de Fourier de una función continua diverge en ciertos puntos (ejemplos de tales series existen), a base de la serie puede restablecerse la propia función en todos los puntos. Veremos, en fin, que desde cierto punto de vista resulta natural considerar la convergencia de las series de Fourier no sólo en el sentido habitual (como convergencia de una sucesión de las sumas parciales en un punto o convergencia uniforme), sino de manera sumamente diferente, a saber, en el sentido de la media cuadrática (véanse los pp. 55.8 y 55.9).

55.2. COEFICIENTES DE FOURIER QUE TIENDEN A CERO

Es de gran importancia en la teoría de las series trigonométricas el hecho de que los coeficientes de Fourier de una función absolutamente integrable tienden hacia cero cuando $n \to \infty$. Se deduce el hecho de una afirmación algo más general que se demuestra abajo y se emplea frecuentemente en las investigaciones concernientes a las series de Fourier y los problemas contiguos.

Teorema 2 (de Riemann). *Si la función f es absolutamente integrable en el intervalo (a, b), sea éste finito o infinito, entonces*

$$\lim_{\nu \to \infty} \int_a^b f(x) \cos \nu x \, dx = \lim_{\nu \to \infty} \int_a^b f(x) \operatorname{sen} \nu x \, dx = 0.$$

Corolario. *Los coeficientes de Fourier* (55.6) *de una función absolutamente integrable en el segmento $[-\pi, \pi]$ tienden hacia cero cuando $n \to \infty$.*

Antes de demostrar estas afirmaciones, introduzcamos algunas nociones que en adelante serán de empleo frecuente.

Definición 4. *Para cualquier función f, definida en todo el eje numérico, la clausura de un conjunto de puntos en los cuales $f(x) \neq 0$ lleva el nombre de portador de la función y se designa* $\operatorname{supp} f$ *).

Definición 5. *Una función f, definida en todo el eje numérico, se denomina finita, siempre que su portador está contenido en cierto segmento finito.*

Definición 6. *Para todo conjunto X dispuesto en una recta numérica, la función*

$$\chi(x) = \chi_X(x) = \begin{cases} 1, & \text{si } x \in X, \\ 0, & \text{si } x \notin X, \end{cases}$$

recibe el nombre de función característica del conjunto X.

---

*) Proviene de la palabra latina supportus.

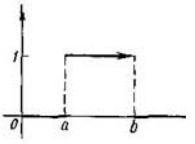


Fig. 230

En la fig. 230 se expone la función característica de un semiintervalo del tipo $[a, b)$.

Definición 7. *Una función f, definida en todo el eje numérico, se llama escalonada finita, siempre que es combinación lineal de un número finito de funciones características de los semiintervalos disjuntos dos a dos $[a_i, b_i)$, $i = 1, 2, \ldots, m$, es decir, si puede ser representada en la forma*

$$f(x) = \sum_{i=1}^{m} \lambda_i \chi_i(x) \tag{55.7}$$

(fig. 231), *donde $\chi_i(x)$ es la función característica del intervalo $[a_i, b_i)$ y λ_i, $i = 1, \ldots, m$, son ciertos números reales.*

No es difícil convencerse de que si no se requiere que los semiintervalos $[a_i, b_i)$, $i = 1, 2, \ldots, m$, sean disjuntos dos a dos, se obtendrá una definición equivalente. Esto se deduce de que la intersección de un número finito de los semiintervalos acotados en consideración es también un semiintervalo del mismo tipo.

Evidentemente, toda función de la forma (55.7) es finita.

Una función escalonada finita f es integrable en todo el eje numérico y, además, si está dada por la fórmula (55.7), entonces

$$\int_{-\infty}^{+\infty} f(x)dx = \sum_{i=1}^{m} \lambda_i \int_{-\infty}^{+\infty} \chi_i(x)dx = \sum_{i=1}^{m} \lambda_i \int_{a_i}^{b_i} dx = \sum_{i=1}^{m} \lambda_i (b_i - a_i).$$

Ejercicio 4. Demuéstrese que cualquier función continua en un segmento es el límite de una sucesión uniformemente convergente de las funciones escalonadas finitas cuyos portadores pertenecen al mismo segmento.

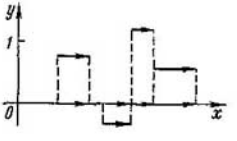


Fig. 231

Lema 2. *Para cualquier función f, absolutamente integrable en un intervalo finito o infinito de extremos a y b, $-\infty \leqslant a < b \leqslant +\infty$, existe una sucesión de tales funciones escalonadas finitas φ_n, $n = 1, 2, \ldots$, que*

1°) $\operatorname{supp} \varphi_n \subset (a, b)$,

2°) $\lim\limits_{n\to\infty} \int\limits_a^b |f(x) - \varphi_n(x)|\,dx = 0$.

DEMOSTRACIÓN. Supongamos que la función f es absolutamente integrable en un intervalo cuyos extremos son a y b. Admitamos, para concretar, que es integrable en cualquier segmento

$$[\xi, \eta], \quad -\infty \leqslant a < \xi < \eta < b \leqslant +\infty$$

(el caso general de una función absolutamente integrable, véase el p. 55.1, se reduce con facilidad al caso que se considera). Entonces, de acuerdo con la definición de integral impropia, para cualquier número fijo $\varepsilon > 0$ existen tales números ξ y η que

$$\int\limits_a^{\xi} |f(x)|\,dx + \int\limits_{\eta}^{b} |f(x)|\,dx < \frac{\varepsilon}{2}. \tag{55.8}$$

La función f es integrable según Riemann en el segmento $[\xi, \eta]$ y, por consiguiente, si designamos con s_τ la suma inferior de Darboux de la función f, correspondiente a cierta partición τ del segmento $[\xi, \eta]$, tendremos

$$\lim_{\delta_\tau \to 0} s_\tau = \int\limits_{\xi}^{\eta} f(x)dx,$$

donde δ_τ es la finura de la partición τ. Por ello, existe una partición $\tau_0 = \{x_i\}_{i=1}^{k}$ del segmento $[\xi, \eta]$ tal que si s_{τ_0} es la suma inferior de Darboux para la función f, correspondiente a la partición τ_0, es decir, si

$$s_{\tau_0} = \sum_{i=1}^{k} m_i \Delta x_i, \quad m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x),$$

$$\Delta x_i = x_i - x_{i-1}, \quad i = 1, 2, \ldots, k,$$

entonces

$$0 \leqslant \int\limits_{\xi}^{\eta} f(x)dx - s_{\tau_0} < \frac{\varepsilon}{2},$$

donde ε es el número fijado anteriormente.

Pongamos

$$\varphi(x) = \begin{cases} m_i, & \text{si } x_{i-1} \leqslant x < x_i, \quad i = 1, 2, \ldots, k, \\ 0, & \text{si } x < \xi \ \text{ó} \ x \geqslant \eta. \end{cases} \tag{55.9}$$

Evidentemente, $\varphi(x)$ es la función escalonada finita,

$$\operatorname{supp}\varphi \subset [\xi, \eta] \subset (a, b) \quad y \quad \int_{\xi}^{\eta} \varphi(x)dx = \sum_{i=1}^{k} m_i \Delta x_i = s_{\tau_0}.$$

Por consiguiente,

$$\int_{\xi}^{\eta} \Big[f(x) - \varphi(x)\Big]dx = \int_{\xi}^{\eta} f(x)dx - \int_{\xi}^{\eta} \varphi(x)dx < \frac{\varepsilon}{2}, \tag{55.10}$$

con la particularidad de que por cuanto $\varphi(x) \leqslant f(x)$, $\xi \leqslant x < \eta$, entonces

$$f(x) - \varphi(x) = |f(x) - \varphi(x)| \geqslant 0.$$

De las desigualdades (55.8) y (55.10) tenemos:

$$\int_{a}^{b} |f(x) - \varphi(x)|dx = \int_{a}^{\xi} |f(x)|dx + \int_{\xi}^{\eta} [f(x) - \varphi(x)]dx + \int_{\eta}^{b} |f(x)|dx < \varepsilon.$$

Suponiendo, por ejemplo, $\varepsilon = \dfrac{1}{n}$ y denotando las correspondientes funciones escalonadas finitas φ mediante φ_n, $n = 1, 2, \ldots$, obtendremos una sucesión de funciones escalonadas finitas φ_n, para la cual se cumple la afirmación del lema. □

DEMOSTRACIÓN DEL TEOREMA. Sea $\chi(x)$ la función característica del semiintervalo $[\xi, \eta)$. Para todo intervalo $(a, b) \supset [\xi, \eta]$ tendremos

$$\lim_{\nu \to \infty} \int_{a}^{b} \chi(x) \operatorname{sen} \nu x dx = \lim_{\nu \to \infty} \int_{\xi}^{\eta} \operatorname{sen} \nu x dx = \lim_{\nu \to \infty} \frac{\cos \nu\xi - \cos \nu\eta}{\nu} = 0,$$

pues

$$\left|\frac{\cos \nu\xi - \cos \nu\eta}{\nu}\right| \leqslant \frac{|\cos \nu\xi| + |\cos \nu\eta|}{\nu} \leqslant \frac{2}{\nu} \to 0, \quad \nu \to \infty.$$

Por cuanto cualquier función escalonada finita es una combinación lineal de un número finito de funciones características de los semiintervalos del tipo considerado, la afirmación del teorema queda válida también para toda función escalonada finita.

Ahora, si la función f es absolutamente integrable sobre el intervalo de extremos a y b, $-\infty \leqslant a < b \leqslant +\infty$, para cualquier número $\varepsilon > 0$, con arreglo al lema, existe una función escalonada finita φ tal que

$$\int_{a}^{b} |f(x) - \varphi(x)|dx < \frac{\varepsilon}{2}.$$

Para esta función escalonada (por cuanto para las funciones escalonadas el teorema ya ha sido demostrado) existe tal ν_ε que con $|\nu| \geqslant \nu_\varepsilon$

$$\left|\int_{a}^{b} \varphi(x) \operatorname{sen} \nu x dx\right| < \frac{\varepsilon}{2}.$$

Por ello, al hacer uso de la identidad $f(x) = [f(x) - \varphi(x)] + \varphi(x)$, para $|\nu| \geqslant \nu_\varepsilon$ obtendremos

$$\left|\int_a^b f(x)\operatorname{sen}\nu x\,dx\right| \leqslant \left|\int_a^b [f(x) - \varphi(x)]\operatorname{sen}\nu x\,dx\right| +$$

$$+ \left|\int_a^b \varphi(x)\operatorname{sen}\nu x\,dx\right| \leqslant \int_a^b \Big|f(x) - \varphi(x)\Big|\,dx +$$

$$+ \left|\int_a^b \varphi(x)\operatorname{sen}\nu x\,dx\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto es indicio de que $\lim\limits_{\nu\to\infty} \int\limits_a^b f(x)\operatorname{sen}\nu x\,dx = 0$.

De modo análogo se demuestra que

$$\lim_{\nu\to\infty} \int_a^b f(x)\cos\nu x\,dx = 0. \ \square$$

55.3. INTEGRAL DE DIRICHLET. PRINCIPIO DE LOCALIZACIÓN

Sea la función $f(x)$ absolutamente integrable en el segmento $[-\pi, \pi]$. Hallemos una expresión, cómoda para investigar, de la suma parcial $S_n(x; f)$ de la serie de Fourier de la función f, llamada también simplemente *suma de Fourier de n-ésimo orden*, $i = 0, 1, 2, \ldots$, *de dicha función*. Al sustituir en $S_n(x; f)$ la expresión para los coeficientes de Fourier (55.6), obtendremos

$$S_n(x; f) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \operatorname{sen} kx =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)dt + \sum_{k=1}^{n} \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)(\cos kt \cos kx + \operatorname{sen} kt \operatorname{sen} kx)\ dt =$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \left[\frac{1}{2} + \sum_{k=1}^{n} \cos k(t - x)\right] dt. \quad (55.11)$$

Pongamos

$$D_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \cos kt, \quad (55.12)$$

entonces la fórmula (55.11) se escribirá en la forma

$$S_n(x; f) = \frac{1}{\pi} \int_{-\pi}^{\pi} D_n(t - x) f(t)dt. \quad (55.13)$$

La función $D_n(t)$ se llama *núcleo de Dirichlet* y la integral que figura en el segundo miembro de la igualdad (55.13), *integral de Dirichlet*.

Lema 3. *El núcleo de Dirichlet*:

1) *es una función par continua* 2π-periódica, con la particularidad de que

$$D_n(0) = n + \frac{1}{2};$$

2) *satisface la condición*

$$\frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)dt = 1; \tag{55.14}$$

3) *cuando* $t \neq 2\pi k$, $k = 0, \pm 1, \pm 2, \ldots$:

$$D_n(t) = \frac{\operatorname{sen}\left(n + \frac{1}{2}\right)t}{2\operatorname{sen}\frac{t}{2}}. \tag{55.15}$$

DEMOSTRACIÓN. La continuidad, la paridad y la existencia de un *período*, igual a 2π, para el núcleo de Dirichlet $D_n(t)$ se desprenden inmediatamente de su definición, es decir, de la fórmula (55.12). De la misma fórmula se infiere también la igualdad (55.14): para obtenerla, es suficiente integrar en el segmento $[-\pi, \pi]$ ambos miembros de la igualdad (55.12):

$$\int_{-\pi}^{\pi} D_n(t)dt = \frac{1}{2}\int_{-\pi}^{\pi} dt + \sum_{k=1}^{n}\int_{-\pi}^{\pi} \cos kt\,dt = \pi,$$

pues, cuando $k = 1, 2, \ldots$: $\int_{-\pi}^{\pi} \cos kt\,dt = 0$.

Demostremos ahora la fórmula (55.15). Tenemos:

$$D_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \cos kt = \frac{1}{2\operatorname{sen}\frac{t}{2}}\left(\operatorname{sen}\frac{t}{2} + \sum_{k=1}^{n} 2\operatorname{sen}\frac{t}{2}\cos kt\right) =$$

$$= \frac{1}{2\operatorname{sen}\frac{t}{2}}\left[\operatorname{sen}\frac{t}{2} + \sum_{k=1}^{n}\left(\operatorname{sen}\frac{2k+1}{2}t - \operatorname{sen}\frac{2k-1}{2}t\right)\right] =$$

$$= \frac{\operatorname{sen}\left(n + \frac{1}{2}\right)t}{2\operatorname{sen}\frac{t}{2}}, \quad t \neq 2\pi k, \quad k = 0, \pm 1, \pm 2, \ldots .$$

Observemos que por ser par el núcleo de Dirichlet,

$$\int_{-\pi}^{0} D_n(t)dt = \int_{0}^{\pi} D_n(t)dt,$$

por lo cual

$$\int_{-\pi}^{\pi} D_n(t)dt = 2\int_{0}^{\pi} D_n(t)dt.$$

De aquí y de la propiedad 2° del núcleo de Dirichlet se deduce que

$$\frac{2}{\pi}\int_{0}^{\pi} D_n(t)dt = 1. \tag{55.16}$$

Indiquemos, además, que el segundo miembro de la igualdad (55.15) tiene sentido sólo cuando $t \neq 2\pi k$, siendo k entero. Pero, como

$$\lim_{t \to 2k\pi} \frac{\operatorname{sen}\left(n + \frac{1}{2}\right)t}{2 \operatorname{sen}\frac{t}{2}} = \lim_{t \to 2k\pi} D_n(t) = n + \frac{1}{2},$$

la función $\dfrac{\operatorname{sen}\left(n + \frac{1}{2}\right)t}{2 \operatorname{sen}\frac{t}{2}}$ puede ser definida adicionalmente para $t = 2\pi k$, $k = 0, \pm 1, \pm 2, \ldots$, considerando que su valor en cada uno de estos puntos es igual, por definición, a $n + \frac{1}{2}$. Una función, definida adicionalmente de este modo, es continua para $t = 2\pi k$, cualquiera que sea k entero.

Volvamos a considerar la función f, absolutamente integrable en el segmento $[-\pi, \pi]$. Para nosotros será de interés, en particular, el límite de la sucesión de sumas parciales $S_n(x)$ de su serie de Fourier. Hemos de notar que el paso directo al límite, para $n \to \infty$, en el segundo miembro de la igualdad (55.13), es decir, el paso al límite bajo el signo de integral, no es posible, puesto que el límite del núcleo de Dirichlet para $n \to \infty$ no existe. Prolongamos la función f desde el semiintervalo $[-\pi, \pi)$ en una función 2π-periódica y designémosla también con f (véase en el p. 55.1 la información más detallada sobre la prolongación periódica).

Demostremos el siguiente lema.

Lema 4. *Para la suma parcial de Fourier $S_n(x; f)$ de una función 2π-periódica absolutamente integrable f son válidas las fórmulas*

$$S_n(x; f) = \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)f(x + t)dt \tag{55.17}$$

y

$$S_n(x;f)=\frac{1}{\pi}\int_0^{\pi} D_n(t)[f(x+t)+f(x-t)]dt. \tag{55.18}$$

Corolario. *Para cualesquiera* $\delta\in(0,\pi)$, $x\in[-\pi,\pi]$ *la suma parcial* $S_n(x;f)$ *de la serie de Fourier de una función* 2π*-periódica absolutamente integrable* f *posee la siguiente representación integral asintótica:*

$$S_n(x;f)=\frac{1}{\pi}\int_0^{\delta} D_n(t)[f(x+t)+f(x-t)]dt+o(1),\quad n\to\infty. \tag{55.19}$$

DEMOSTRACIÓN DEL LEMA. Realicemos en la integral de Dirichlet (55.13) el cambio de variable de integración $u=t-x$:

$$S_n(x;f)=\frac{1}{\pi}\int_{-x}^{\pi} D_n(t-x)f(t)dt=$$

$$=\frac{1}{\pi}\int_{-\pi-x}^{\pi-x} D_n(u)f(x+u)du=\frac{1}{\pi}\int_{-\pi}^{\pi} D_n(u)f(x+u)du. \tag{55.20}$$

Aquí hemos aprovechado otra vez la circunstancia de que la integral de una función periódica extendida al segmento cuya longitud es igual al período de la función, no depende de la posición de dicho segmento en el eje real (véase el p. 55.1) y hemos aplicado esta propiedad a la función $D_n(u)f(x+u)$ que es 2π-periódica según u. Así pues, la fórmula (55.17) queda demostrada.

Con el objeto de demostrar la fórmula (55.18), representemos el segundo miembro de la igualdad (55.20) en forma de una suma de dos integrales cuyos intervalos de integración son $[-\pi,0]$ y $[0,\pi]$; en la primera integral realicemos el cambio de variable $u=-t$ y hagamos uso de que el núcleo de Dirichlet es par:

$$D_n(-u)=D_n(u)$$

(véase el lema 3). Como resultado tendremos:

$$S_n(x;f)=\frac{1}{\pi}\int_{-\pi}^{0} D_n(u)f(x+u)du=$$

$$=\frac{1}{\pi}\int_{-\pi}^{0} D_n(u)f(x+u)du+\frac{1}{\pi}\int_0^{\pi} D_n(u)f(x+u)du=$$

$$=\frac{1}{\pi}\int_0^{\pi} D_n(t)f(x-t)dt+\frac{1}{\pi}\int_0^{\pi} D_n(u)f(x+u)du=$$

$$= \frac{1}{\pi} \int_0^{\pi} D_n(t)[f(x + t) + f(x - t)]dt.$$

La fórmula (55.18) también queda demostrada. □

DEMOSTRACIÓN DEL COROLARIO. Fijemos un número δ, $0 < \delta < \pi$, y representemos el segundo miembro de (55.18) como una suma de dos integrales de la manera siguiente:

$$S_n(x; f) = \frac{1}{\pi} \int_0^{\delta} + \frac{1}{\pi} \int_{\delta}^{\pi}. \tag{55.21}$$

Por cuanto la función $\dfrac{1}{2 \operatorname{sen} \dfrac{t}{2}}$ es continua y, por lo tanto, acotada en el segmento $[\delta, \pi]$ (a saber, para todo $t \in [\delta, \pi]$: $0 < \dfrac{1}{2 \operatorname{sen} \dfrac{t}{2}} \leqslant \dfrac{1}{2 \operatorname{sen} \dfrac{\delta}{2}}$), mientras la función $f(x + t) + f(x - t)$, para cualquier $x \in [-\pi, \pi]$ fijo es 2π-periódica según t y absolutamente integrable en el segmento $[-\pi, \pi]$, entonces en $[\delta, \pi]$ será absolutamente integrable también el producto de dichas funciones $\dfrac{f(x + t) + f(x - t)}{2 \operatorname{sen} \dfrac{t}{2}}$. Por esta razón, de acuerdo con el teorema de Riemann (véase el teorema 2 en el p. 55.2), la segunda integral en el segundo miembro de la igualdad (55.21) tiende hacia cero cuando $n \to \infty$, es decir,

$$\frac{1}{\pi} \int_{\delta}^{\pi} \frac{f(x + t) + f(x - t)}{2 \operatorname{sen} \frac{t}{2}} \operatorname{sen} \left(n + \frac{1}{2}\right) t dt = o(1), \quad n \to \infty.$$

Sustituyendo esta expresión en la fórmula (55.21), obtendremos la (55.19). □

De la fórmula (55.19) se infiere una propiedad importante de las series de Fourier, llamada principio de localización. Enunciémoslo en forma de un teorema.

Teorema 3 (principio de localización). *Si f es una función 2π-periódica absolutamente integrable, la existencia y el valor del límite de la sucesión de sus sumas parciales de Fourier $S_n(x; f)$ en todo punto $x_0 \in [-\pi, \pi]$ depende sólo de la existencia y del valor del límite, para $n \to \infty$ de la integral*

$$\frac{1}{\pi} \int_0^{\delta} D_n(t)[f(x_0 + t) + f(x_0 - t)]dt,$$

donde δ es un número positivo tan pequeño como se quiera.

Subrayemos que en la expresión subintegral de la integral citada figuran sólo los valores de la función f en el segmento $[x_0 - \delta, x_0 + \delta]$ y, de este modo, la existen-

cia y el valor del límite de las sumas parciales de la serie de Fourier de la función f sólo dependen de sus propiedades en el entorno del punto x_0, o, como suele decirse, de sus propiedades locales en las cercanías del punto x_0.

Del principio de localización proviene que si en un entorno, tan pequeño como se quiera, del punto x_0 las funciones f y g coinciden, entonces los límites $\lim_{n\to\infty} S(x_0; f)$ y $\lim_{n\to\infty} S(x_0; g)$ existen o no simultáneamente con la particularidad de que si dichos límites existen, son iguales. Es de interés particular que las series de Fourier de tales funciones son, en el caso general, diferentes, pues en las fórmulas para los coeficientes de Fourier figuran los valores de la función en todo el segmento $[-\pi, \pi]$.

55.4. CONVERGENCIA DE LA SERIE DE FOURIER EN UN PUNTO

En este punto se considerarán las funciones 2π-periódicas absolutamente integrables en un segmento de longitud 2π, las cuales tienen sólo puntos de discontinuidad de primera especie, a consecuencia de lo cual en todo punto x_0 del eje numérico existen límites unilaterales:

$$\lim_{h\to+0} f(x_0 + h) = f(x_0 + 0), \quad \lim_{h\to+0} f(x_0 - h) = f(x_0 - 0).$$

Definición 8 (de Lebesgue [*]). *Un punto x_0 se llama punto regular de la función f, si*

$$f(x_0) = \frac{f(x_0 + 0) + f(x_0 - 0)}{2}.$$

Evidentemente, todo punto de continuidad de una función es su punto regular.

Si x_0 es un punto de discontinuidad de primera especie, a título de sus derivadas unilaterales $f'_+(x)$ y $f'_-(x)$ intervendrán los límites

$$f'_+(x) = \lim_{h\to+0} \frac{f(x + h) - f(x + 0)}{h},$$

$$f'_-(x) = \lim_{h\to+0} \frac{f(x - h) - f(x - 0)}{-h}.$$

En aquel caso cuando la función es continua en el punto x, y, por ende, $f(x + 0) = f(x - 0) = f(x)$, la definición enunciada de las derivadas unilaterales coincide con la aducida anteriormente (véase el p. 9.1).

Para que la enunciación del teorema sobre la convergencia de la serie de Fourier sea más cómoda, introduzcamos una designación

$$f_x^*(t) \stackrel{\text{def}}{=} f(x + t) + f(x - t) - f(x + 0) - f(x - 0). \tag{55.22}$$

Es obvio que en el punto regular x la función $f_x^*(t)$ tiene por expresión

$$f_x^*(t) = f(x + t) + f(x - t) - 2f(x).$$

---

[*] H. L. Lebesgue (1875—1941), matemático francés.

Nos hará falta el siguiente lema sencillo.

Lema 5. *Para la función f, 2π-periódica y absolutamente integrable en el segmento de longitud 2π, las integrales*

$$\int_0^{\delta} \frac{|f_x^*(t)|}{t}\,dt, \quad 0 < \delta \leqslant \pi, \quad y \int_{\delta}^{\pi} \frac{|f_x^*(t)|}{2\,\operatorname{sen}\frac{t}{2}}\,dt \tag{55.23}$$

convergen o divergen simultáneamente.

DEMOSTRACIÓN. En efecto, para cualquier δ, $0 < \delta < \pi$ la función $\dfrac{1}{2\,\operatorname{sen}\frac{t}{2}}$ es continua y, por ende, integrable según Riemann en el segmento $[\delta, \pi]$. Entre tanto, la función $f_x^*(t)$ (x es fijo) es absolutamente integrable en dicho segmento y, por consiguiente, el producto $\dfrac{f_x^*(t)}{2\,\operatorname{sen}\frac{t}{2}}$ es también absolutamente integrable en el segmento $[\delta, \pi]$, es decir, para cualquier δ, $0 < \delta < \pi$, la integral

$$\int_{\delta}^{\pi} \frac{|f_x^*(t)|}{2\,\operatorname{sen}\frac{t}{2}}\,dt \tag{55.24}$$

converge (véase el lema 2 en el p. 33.5).

Elijamos ahora un $\delta > 0$ de modo tal que en el segmento $[0, \delta]$ la función $f_x^*(t)$ no tenga puntos singulares (véase el p. 55.1), a excepción, quizás, del punto $t = 0$, es decir, que para todo ε, $0 < \varepsilon < \delta$, sea integrable según Riemann en el segmento $[\varepsilon, \delta]$; esto es siempre factible, puesto que la supuesta integrabilidad absoluta de la función f estipula que f y, por consiguiente, también f_x^* sólo disponen de un número finito de puntos singulares (véase de nuevo el p. 55.1).

Ahora diremos que las funciones $\dfrac{f_x^*(t)}{t}$ y $\dfrac{f_x^*(t)}{2\,\operatorname{sen}\frac{t}{2}}$ son equivalentes para $t \to 0$, pues

$$\lim_{t \to 0} \frac{2\,\operatorname{sen}\frac{t}{2}}{t} = 1;$$

por lo cual, de acuerdo con el criterio de convergencia de las integrales, llamado criterio de comparación (véase el corolario del teorema 1 en el p. 33.3), aplicado a las magnitudes absolutas de las funciones en consideración, las integrales

$$\int_0^{\delta} \frac{|f_x^*(t)|}{t}\,dt, \int_{\delta}^{\delta} \frac{|f_x^*(t)|}{2\,\operatorname{sen}\frac{t}{2}}\,dt$$

convergen o divergen simultáneamente. Siendo convergente la integral (55.24), de aquí se deduce de inmediato que las integrales (55.23) también convergerán o divergerán simultáneamente. □

Teorema 4 (criterio de Dini). *Supongamos que una función 2π-periódica f es absolutamente integrable en un segmento de longitud 2π.*

En este caso, si x es punto de continuidad o punto de discontinuidad de primera especie y, para cierto δ, $0 < \delta < \pi$, la integral

$$\int_0^{\delta} \left| \frac{f_x^*(t)}{t} \right| dt \tag{55.25}$$

converge, entonces la serie de Fourier de la función f converge en el punto x hacia el valor

$$\frac{f(x+0) + f(x-0)}{2}. \tag{55.26}$$

Corolario 1. *Cumplidas las condiciones del teorema, en cualquier punto regular de la función f (en particular, en todos los puntos de continuidad de f) la serie de Fourier de dicha función converge hacia su valor en el punto que se considera.*

Corolario 2. *Si f es una función 2π-periódica absolutamente integrable en un segmento de longitud 2π y si en el punto x existen $f(x+0)$, $f(x-0)$, $f'_+(x)$ y $f'_-(x)$, entonces la serie de Fourier de la función converge en el punto dado hacia el valor* (55.26).

Corolario 3. *La serie de Fourier de una función f, continuamente derivable a trozos en el segmento $[-\pi, \pi]$, converge en todo punto del intervalo $(-\pi, \pi)$ hacia el valor* (55.26), *y en los puntos $x = -\pi$ y $x = \pi$, al valor*

$$\frac{f(-\pi+0) + f(\pi-0)}{2}. \tag{55.27}$$

Corolario 4. *La serie de Fourier de una función continua derivable a trozos en el segmento $[-\pi, \pi]$ converge en todo punto del intervalo $(-\pi, \pi)$ hacia el valor de la función en este punto y en los puntos $x = -\pi$ y $x = \pi$, al valor* (55.27).

DEMOSTRACIÓN DEL TEOREMA. Al utilizar las fórmulas (55.18) y (55.16), tendremos

$$S_n(x; f) - \frac{f(x+0) + f(x-0)}{2} =$$

$$= \frac{1}{\pi} \int_{10}^{\pi} D_n(t)[f(x+t) + f(x-t)]dt - \frac{f(x+0) + f(x-0)}{2} \frac{2}{\pi} \int_0^{\pi} D_n(t)dt =$$

$$= \frac{1}{\pi} \int_0^{\pi} D_n(t)\{[f(x+t) + f(x-t) - f(x+0) - f(x-0)]dt =$$

$$= \frac{1}{\pi} \int_0^{\delta} \frac{f_x^*(t)}{2 \operatorname{sen} \frac{t}{2}} \operatorname{sen}\left(n + \frac{1}{2}\right)t\,dt. \tag{55.28}$$

Supongamos que la integral (55.25) es convergente. En este caso, según el lema 5, converge también la integral

$$\int_0^{\pi} \frac{|f_x^*(t)|}{2 \operatorname{sen} \frac{t}{2}}\, dt,$$

en otras palabras, la función $\dfrac{f_x^*(t)}{2 \operatorname{sen} \frac{t}{2}}$ es absolutamente integrable en el segmento $[0, \pi]$. Por eso, de acuerdo con el teorema de Riemann (véase el p. 55.2),

$$\lim_{n \to \infty} \frac{1}{\pi} \int_0^{\pi} \frac{f_x^*(t)}{2 \operatorname{sen} \frac{t}{2}} \operatorname{sen} \left(n + \frac{1}{2}\right) t\, dt = 0,$$

por consiguiente, en virtud de (55.28):

$$\lim_{n \to \infty} S_n(x; f) = \frac{f(x + 0) + f(x - 0)}{2}. \quad \square$$

El corolario 1 proviene inmediatamente del teorema, en virtud de la definición de punto regular de una función. $\square$

Demostraremos el corolario 2. De conformidad con el teorema 4, basta probar que si existen los límites $f(x + 0), f(x - 0)$ y las derivadas unilaterales $f'_+(x), f'_-(x)$, la integral (55.25) converge para cierto $\delta > 0$. Ante todo, debido a la existencia del límite finito

$$\lim_{t \to 0} \frac{f_x^*(t)}{t} = \lim_{t \to 0} \left[\frac{f(x + t) - f(x + 0)}{t} + \right.$$

$$\left. + \frac{f(x - t) - f(x - 0)}{t} \right] = f'_+(x) - f'_-(x),$$

la función $\dfrac{f_x^*(t)}{t}$ está acotada en cierto entorno del punto $t = 0$. Por eso existe tal δ, $0 < \delta < \pi$, que en el segmento $[0, \delta]$ la función $\dfrac{f_x^*(t)}{t}$ está acotada y, por lo tanto, no tiene puntos singulares, a consecuencia de lo cual es integrable según Riemann en este segmento (véase el p. 33.1, como también la observación 4 en el p. 44.7). Una función, integrable según Riemann es integrable absolutamente, razón por la cual la integral (55.25) es finita. $\square$

Con el fin de demostrar el corolario 3, prolongamos la función f, definida en el segmento $[-\pi, \pi]$, de una manera periódica con el período igual a 2π desde el semiintervalo $[-\pi, \pi)$ a todo el eje numérico y designemos la función obtenida con $\bar{f}$. Por definición de la derivabilidad a trozos (véase la definición 1 en el p. 30.2), la función $\bar{f}$ satisface las condiciones del corolario 2. De conformidad con este corolario, la serie de Fourier de la función $\bar{f}$, que coincide, evidentemente, con la serie de

Fourier para f, converge en todo punto x hacia

$$\frac{\bar{f}(x+0)+\bar{f}(x-0)}{2}.$$

Si $x \in (-\pi, \pi)$, se tiene $\bar{f}(x \pm 0) = f(x \pm 0)$ y, por lo tanto,

$\frac{\bar{f}(x+0)+\bar{f}(x-0)}{2} = \frac{f(x+0)+f(x-0)}{2}$. Cuando $x = -\pi$, la serie mencionada converge hacia $\frac{\bar{f}(-\pi+0)+\bar{f}(-\pi-0)}{2}$, y si $x = \pi$, hacia el valor $\frac{\bar{f}(\pi+0)+\bar{f}(\pi-0)}{2}$. Dado que $\bar{f}$ es periódica, se tiene

$$\bar{f}(-\pi-0) = \bar{f}(\pi-0) = f(\pi-0), \bar{f}(\pi+0) = \bar{f}(-\pi+0) = f(-\pi+0).$$

Por esto

$$\frac{\bar{f}(-\pi+0)+\bar{f}(-\pi-0)}{2} = \frac{\bar{f}(\pi+0)+\bar{f}(\pi-0)}{2} = \frac{f(-\pi+0)+f(\pi-0)}{2}. \square$$

El corolario 4 proviene directamente de los corolarios 1 y 3. $\square$

Ha de ser notado que en las fórmulas (55.26) y (55.27) la suma de la serie de Fourier de la función f viene expresada en términos de la propia función f, definida en el segmento $[-\pi, \pi]$, y no mediante su prolongación periódica $\bar{f}$ a todo el eje numérico.

Si la función f satisface las condiciones del corolario 4, es decir, es continua y derivable a trozos en el segmento $[-\pi, \pi]$ y, además, $f(-\pi) = f(\pi)$ (es decir, su prolongación periódica a todo el eje numérico coincide con ella en todo punto de $[-\pi, \pi]$, incluidos los extremos), entonces en todo el segmento $[-\pi, \pi]$ la función f es igual a la suma de su serie de Fourier:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

Por esta razón tal función en todo punto del segmento $[-\pi, \pi]$ puede ser representada con cualquier grado de precisión por la suma parcial de su serie de Fourier, es decir, por una combinación lineal de senos y cosenos de los arcos múltiples (suele decirse también que la función citada se aproxima por la suma de armónicos sencillos *). El hecho de que en el caso dado el período es igual precisamente a 2π no es esencial: el caso en que el período arbitrario $T > 0$ se reduce con facilidad al considerado, simplemente por cambio de variable (véase el p. 55.12).

Ejemplos. 1. Hallemos la serie de Fourier de la función $\operatorname{ch} x$, $-\pi \leqslant x \leqslant \pi$.

*) Se llama armónico sencillo (en la física, principalmente) una expresión de la forma $A \cos nx + B \operatorname{sen} nx$, donde A y B son unas constantes, n es un número natural.

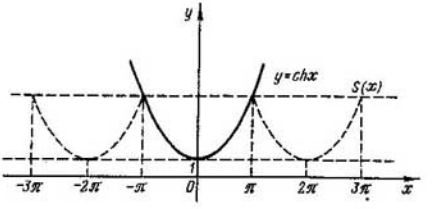


Fig. 232

Calculemos sus coeficientes de Fourier:

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{ch} x\,dx = \frac{\operatorname{sh} x}{\pi}\Big|_{-\pi}^{\pi} = \frac{2\operatorname{sh}\pi}{\pi},$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{ch} x \cos nx\,dx = (-1)^n \frac{2\operatorname{sh}\pi}{\pi(1+n^2)}, \quad n = 1, 2, \ldots .$$

De lo que la función $\operatorname{ch} x$ es par se deduce que para ella $b_n = 0$, $n = 1, 2, \ldots$. La función $\operatorname{ch} x$ es continuamente derivable y, por lo tanto, satisface las condiciones del corolario 4 del teorema 4; además, la función toma valores iguales en los extremos del segmento $[-\pi, \pi]$, por lo cual su serie de Fourier en todos los puntos del segmento $[-\pi, \pi]$ converge hacia la propia función $\operatorname{ch} x$:

$$\operatorname{ch} x = \frac{\operatorname{sh}\pi}{\pi}\left(1 + 2\sum_{n=1}^{\infty} (-1)^n \frac{1}{1+n^2}\cos nx\right).$$

Esta serie converge uniformemente, lo que se desprende de su comparación con la serie numérica convergente $\sum_{n=0}^{\infty} \frac{1}{1+n^2}$.

Las gráficas de la función $\operatorname{ch} x$ y de la suma $S(x)$ de la serie de Fourier correspondiente se exponen en la fig. 232.

2. Hallemos la serie de Fourier de la función $\operatorname{sh} x$, $-\pi \leqslant x \leqslant \pi$.

En virtud de que esta función es impar, tenemos

$$a_n = 0, \quad n = 0, 1, 2, \ldots ,$$

y, luego,

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \operatorname{sh} x \operatorname{sen} nx\,dx = (-1)^{n-1}\frac{2n\operatorname{sh}\pi}{\pi(1+n^2)}, \quad n = 1, 2, \ldots .$$

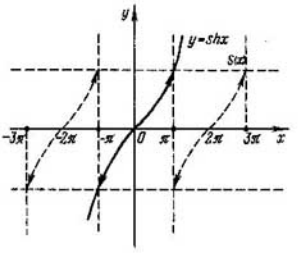


Fig. 233

La función $\operatorname{sh} x$ es continuamente derivable y satisface las condiciones del corolario 4 del teorema 4, pero $\operatorname{sh}(-\pi) \neq \operatorname{sh}\pi$; por eso, en todos los puntos del intervalo $(-\pi, \pi)$ la serie de Fourier de la función $\operatorname{sh} x$ converge hacia la propia función:

$$\operatorname{sh} x = \frac{2 \operatorname{sh}\pi}{\pi} \cdot \sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{1+n^2} \operatorname{sen} nx, \quad -\pi < x < \pi$$

y en los puntos $x = -\pi$ y $x = \pi$, hacia el valor $\dfrac{\operatorname{sh}(-\pi) + \operatorname{sh}\pi}{2} = 0$.

La serie de Fourier de la función $\operatorname{sh} x$ ya no converge uniformemente hacia la función en todo el segmento $[-\pi, \pi]$ (en efecto, de lo contrario su suma debería ser continua en el segmento $[-\pi, \pi]$, mientras que ella tiene discontinuidades en sus extremos). Las gráficas de las funciones $\operatorname{sh} x$ y de la suma $S(x)$ de su serie de Fourier están expuestas, para la comparación, en la fig. 233.

3. Desarrollemos la función

$$f(x) = \frac{\pi - x}{2}, \quad 0 \leqslant x \leqslant 2\pi,$$

en la serie de Fourier.

Aunque la función f parece ser algo artificial, su serie de Fourier es de la forma muy sencilla y permite obtener toda una serie de fórmulas interesantes.

Prolongamos, de manera 2π-periódica, la función f desde el semiintervalo $[0, 2\pi)$ a todo el eje numérico. De resultas se obtendrá una función impar, en virtud de lo cual todos sus coeficientes de Fourier a_n serán nulos; $a_n = 0$, $n = 0, 1, 2, \ldots$.

Calculemos los coeficientes b_n. Integrando por partes, obtendremos

$$b_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - x}{2} \operatorname{sen} nx\, dx =$$

$$= -\frac{1}{2\pi} (\pi - x) \frac{\cos nx}{n} \Bigg|_0^{2\pi} - \frac{1}{2n\pi} \int_0^{2\pi} \cos nx\, dx = \frac{1}{n}.$$

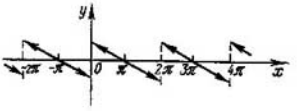


Fig. 234

Así pues,

$$\frac{\pi - x}{2} \sim \sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n}. \tag{55.29}$$

En virtud del corolario 4 del teorema 4, para $0 < x < 2\pi$ tiene lugar la igualdad

$$\frac{\pi - x}{2} = \sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n}. \tag{55.30}$$

Cuando $x = 0$, esta igualdad, evidentemente, no se verifica, puesto que la suma de la serie obtenida para $x = 0$ es nula y $f(0) \neq 0$.

La gráfica de la suma de la serie (55.29) va expuesta en la fig. 234. Diremos que esta serie a ciencia cierta no converge uniformemente en el segmento $[0, 2\pi]$, puesto que su suma no es en éste una función continua (la convergencia uniforme de la serie (55.29) se ha investigado en el p. 36.3).

Al sustituir en (55.30) x por $2x$ y dividir ambos miembros de la igualdad obtenida por 2, obtendremos

$$\frac{\pi}{4} - \frac{x}{2} = \sum_{k=1}^{\infty} \frac{\operatorname{sen} 2kx}{2k}, \quad 0 < x < \pi. \tag{55.31}$$

Sustrayamos esta igualdad de (55.30):

$$\frac{\pi}{4} = \sum_{k=1}^{\infty} \frac{\operatorname{sen} (2k-1)x}{2k-1}, \quad 0 < x < \pi. \tag{55.32}$$

Al sustituir la expresión obtenida para $\frac{\pi}{4}$ en (55.31), obtenemos

$$x = 2 \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\operatorname{sen} nx}{n}.$$

Esta igualdad es justa ya para $x = 0$, y, en virtud de que ambos miembros de la igualdad son impares, también para $-\pi < x < 0$, es decir, en todo el intervalo $(-\pi, \pi)$, pero, por supuesto, no en sus extremos, donde la suma de la serie es igual a cero.

Señalemos, además, que al poner en (55.32) $x = \frac{\pi}{2}$, obtendremos la así llamada serie de Leibniz

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots,$$

con la gue ya hemos chocado antes (véase el p. 37.7, ejemplo 2).

Ejercicios. 5. Hállese la serie de Fourier de la función $f(x) = |x|$, $-\pi \leqslant x \leqslant \pi$, y con ayuda de esta serie calcúlese la suma de la serie $\sum_{k=1}^{\infty} \frac{1}{(2k-1)^2}$.

6. Hállese la serie de Fourier para las funciones

a) $f(x) = x^2$, $0 \leqslant x \leqslant 2\pi$;

b) $f(x) = x^2$, $-\pi \leqslant x \leqslant \pi$;

c) $f(x) = x^2$ para $0 \leqslant x \leqslant \pi$, y al semiintervalo $[-\pi, 0)$ la función f se ha prolongado de manera impar.

Con ayuda de los desarrollos obtenidos calcúlense las sumas de las series

$$\sum_{n=1}^{\infty} \frac{1}{n^2}, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2}.$$

55.5*. CONVERGENCIA DE LAS SERIES DE FOURIER PARA FUNCIONES QUE SATISFACEN LA CONDICIÓN DE HÖLDER

En este punto daremos a conocer una condición suficiente más débil, en comparación con la de derivabilidad unilateral (véase el corolario 2 del teorema 4 en el p. 55.4), que también asegura la convergencia de la integral (55.25) y, por lo tanto, la convergencia de la serie de Fourier de la función f hacia el valor (55.26), siendo f 2π-periódica absolutamente integrable en el segmento de longitud igual al período.

Definición 9. *Una función f, definida en el intervalo (x_0, b), lleva el nombre de función que satisface a la derecha la condición de Hölder de exponente α en el punto x_0, siempre que existen el límite finito a la derecha $f(x_0 + 0)$ y unas constantes $M > 0$ y $\delta > 0$ tales que para cualquier h, que satisfaga la condición $0 < h < \delta$, se verifica la desigualdad*

$$|f(x_0 + h) - f(x_0 + 0)| < Mh^{\alpha}. \tag{55.33}$$

Una función f, definida en el intervalo (a, x_0) se llama función que satisface a la izquierda la condición de Hölder de exponente α en el punto x_0, siempre que existen el límite finito a la izquierda $f(x_0 - 0)$, y unas constantes $M > 0$ y $\delta > 0$ tales que para cualquier h que satisfaga la condición $0 < h < \delta$, se verifica la desigualdad

$$|f(x_0 - h) - f(x_0 - 0)| < Mh^{\alpha}. \tag{55.34}$$

Una función f, que satisface en el punto x_0 la condición de Hölder de cierto exponente tanto a la derecha como a la izquierda, se denomina función que satisface la condición de Hölder de exponente dado en el punto que se considera.

Una función, definida en cierto segmento, se llama función que satisface la condición de Hölder de exponente dado en dicho segmento, si en todo punto de éste satisface la condición de Holder de exponente citado, y además, en todo punto interior del segmento tanto a la derecha como a la izquierda; en el extremo izquierdo del segmento a la derecha, y en el extremo derecho, a la izquierda.

Observemos que la así llamada condición clásica de Hölder de exponente dado consiste en lo siguiente: *de la función f se dice que ella satisface en el punto x la condición clásica de Hölder de exponente $\alpha > 0$, si existen tales $\delta > 0$ y $M > 0$ que para cualesquiera h, $|h| < \delta$, se verifica la desigualdad*

$$|f(x + h) - f(x)| \leqslant M|h|^{\alpha}.$$

Es obvio que en este caso, gracias a la condición $\alpha > 0$, la función f es siempre continua en el punto x: de $\lim\limits_{h \to 0} |h|^{\alpha} = 0$, $\alpha > 0$, se deduce que $\lim\limits_{h \to 0} f(x + h) = f(x)$.

Análogamente se determinan las condiciones clásicas unilaterales de Hölder.

Así pues, la diferencia de la condición de Hölder en consideración de la clásica consiste, particularmente, en que, de conformidad con nuestra definición, una función que satisface la condición de Hölder en un punto puede ser discontinua en el mismo.

La condición de Hölder de exponente uno se llama, comúnmente, *condición de Lipschitz* *).

Ejercicios. 7. Demuéstrese que si una función satisface en cierto punto la condición de Hölder de exponente α, entonces para $0 < \beta < \alpha$ satisface también en el mismo punto la condición de Hölder de exponente β.

8. Demuéstrese que si una función tiene en cierto segmento una derivada acotada, satisface en el mismo la condición de Lipschitz con una misma constante M.

9. Demuéstrese que si una función satisface en cierto segmento la condición clásica de Hölder de exponente $\alpha > 1$, es en este segmento constante.

10. Demuéstrese que la función $f(x) = x^{\alpha}$, $x \geqslant 0$, $0 < \alpha \leqslant 1$, satisface en el punto $x = 0$ la condición de Hölder de exponente α, y no satisface en el mismo ninguna condición de Hölder de exponente $\beta > \alpha$.

Teorema 5. *Supongamos que la función f es absolutamente integrable en el segmento $[-\pi, \pi]$. Si satisface en el punto $x \in (-\pi, \pi)$ la condición de Hölder de exponente α, $\alpha > 0$, entonces su serie de Fourier converge en este punto y su suma es igual a*

$$\frac{f(x + 0) + f(x - 0)}{2};$$

en particular, si la función, además, es continua en el punto $x \in (-\pi, \pi)$, su serie de Fourier converge al valor de la función en el mismo punto:

$$\lim_{n \to \infty} S_n(x; f) = f(x).$$

*) R. Lipschitz (1832 — 1903), matemático alemán.

Si la función f satisface la condición de Hölder a la derecha en el punto $x = -\pi$, y a la izquierda en el punto $x = \pi$, entonces su serie de Fourier converge en estos puntos y su suma en los mismos es igual a

$$\frac{f(-\pi) + f(\pi)}{2}.$$

DEMOSTRACIÓN. Elijamos δ, $0 < \delta < \pi$, de modo tal que, primero, en el segmento $[0, \delta]$ no haya otros puntos singulares de la función $\frac{f_x^*(t)}{t^\alpha}$, a excepción, quizás, del punto $x = 0$, y, segundo, que para cualesquiera h, $|h| < \delta$, la función f satisfaga las condiciones de Hölder (55.33) y (55.34) en el punto x. Entonces, en vista de la fórmula (55.22), para la función $f_x^*(t)$, tendremos

$$\left|\frac{f_x^*(t)}{t}\right| \leqslant \left|\frac{f(x+t) - f(x+0)}{t}\right| + \left|\frac{f(x-t) - f(x-0)}{t}\right| \leqslant \frac{2M}{t^{1-\alpha}}.$$

Por cuanto la integral $\int_0^\delta \frac{dt}{t^{1-\alpha}}$, $\alpha > 0$, converge, entonces, en virtud del criterio de comparación, converge también en nuestro caso la integral (55.25). Por ello, el teorema 5 proviene del teorema 4. □

Como conclusión observemos que si la función f tiene en el punto x una derivada derecha f'_+, entonces f satisface en este punto a la derecha la condición de Hölder de exponente 1. Efectivamente, del hecho de que el límite finito

$$\lim_{h \to 0} \frac{f(x+h) - f(x+0)}{h} = f'_+(x)$$

existe, se deduce que se encontrará tal $\delta > 0$ que para cualesquiera h, $|h| < \delta$, se verificará la desigualdad

$$\left|\frac{f(x+h) - f(x+0)}{h} - f'_+(x)\right| < 1,$$

de donde, al poner $M \overset{\text{def}}{=} |f'_+(x)| + 1$, obtendremos

$$-M < \frac{f(x+h) - f(x+0)}{h} < M;$$

por consiguiente,

$$|f(x+h) - f(x+0)| \leqslant M|h|, \quad |h| < \delta.$$

La afirmación análoga es lícita también para las derivadas a la izquierda.

Problema 35. Una función f, definida en el segmento $[a, b]$, se llama *función de la clase de Hölder $H^\alpha(M)$* en dicho segmento, si para cualquier par de puntos x y $x + h$ del mismo, $x \in [a, b]$, $x + h \in [a, b]$, se cumple la desigualdad

$$|f(x+h) - f(x)| \leqslant M|h|^\alpha,$$

en otras palabras, si la función f satisface la condición clásica de Hölder de un mismo exponente α y con una misma constante M en todos los puntos del segmento $[a, b]$.

Demuéstrese que si una función 2π-periódica y absolutamente integrable en un segmento de longitud 2π pertenece en cierto segmento $[a, b]$ a la clase de Hölder $H^{\alpha}(M)$, $0 < \alpha \leqslant 1$, $M > 0$, entonces en todo segmento $[a', b']$, contenido en el intervalo (a, b): $0 < a < a' < b' < b < 2\pi$, la serie de Fourier de la función f converge hacia esta función uniformemente.

55.6. SUMACIÓN DE LAS SERIES DE FOURIER POR EL MÉTODO DE MEDIAS ARITMÉTICAS

Supongamos que la función f es absolutamente integrable en el segmento $[-\pi, \pi]$ y satisface la condición $f(-\pi) = f(\pi)$, y, por lo tanto, es prolongable de modo 2π-periódico a todo el eje real. Sean $S_n(x)$ sus sumas de Fourier y $D_n(x)$, los núcleos de Dirichlet, $n = 0, 1, 2, \ldots$ (véanse (55.11) y (55.12)).

Consideraremos las medias aritméticas:

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \ldots + S_n(x)}{n+1},$$

$$\Phi_n(x) = \frac{D_0(x) + D_1(x) + \ldots + D_n(x)}{n+1}, \quad n = 0, 1, 2, \ldots. \tag{55.35}$$

La suma $\sigma_n(x)$ se llama *suma de Fejer* *) de n-ésimo orden de la función f, mientras que $\Phi_n(x)$ es el *núcleo de Fejer* de n-ésimo orden.

De la fórmula

$$S_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(u) f(x+u)\, du$$

(véase (55.17)), obtenemos

$$\sigma_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_n(u) f(x+u)\, du. \tag{55.36}$$

Analicemos el comportamiento de las sumas $\sigma_n(x)$ cuando $n \to \infty$, es decir, examinemos la sumación de la serie de Fourier por el *método de medias aritméticas* (véase el p. 35.15).

Estudiaremos ante todo las propiedades del núcleo de Fejer.

Lema 6. *El núcleo de Fejer posee las siguientes propiedades.*

1°. *La función $\Phi_n(x)$ es par, continua y 2π-periódica, con la particularidad de que*

$$\Phi_n(0) = \frac{n+1}{2}; \tag{55.37}$$

2°. *Para todo t el núcleo de Fejer es no negativo:* $\Phi_n(t) \geqslant 0$;

$$3°. \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_n(t)\, dt = \frac{2}{\pi} \int_{0}^{\pi} \Phi_n(t)\, dt = 1;$$

*) L. Fejer (1880 — 1959), matemático húngaro.

4°. *Cuando* $t \neq 2k\pi$, $k = 0, \pm 1, \pm 2, \ldots$:

$$\Phi_n(t) = \frac{\operatorname{sen}^2 \frac{n+1}{2} t}{2(n+1)\operatorname{sen}^2 \frac{t}{2}}.$$

Corolario. *Para todo* δ *fijo,* $0 < \delta < \pi$, *tiene lugar la igualdad*

$$\lim_{n \to \infty} \max_{\delta \leqslant |t| \leqslant \pi} \Phi_n(t) = 0. \tag{55.38}$$

DEMOSTRACIÓN. Demostremos al principio la propiedad 1°. La paridad, la continuidad y el carácter periódico del núcleo de Fejer provienen inmediatamente, en virtud de la fórmula (55.35), de las mismas propiedades del núcleo de Dirichlet (véase el lema 3 en el p. 55.3). Luego, por cuanto $D_k(0) = k + \frac{1}{2}$, $k = 0, 1, 2, \ldots$, entonces

$$\Phi_n(0) = \frac{1}{n+1} \sum_{k=0}^{n} D_k(0) = \frac{1}{n+1} \sum_{k=0}^{n} \left(k + \frac{1}{2}\right) =$$

$$= \frac{1}{n+1} \left[\frac{n(n+1)}{2} + \frac{n+1}{2}\right] = \frac{n+1}{2}.$$

Demostremos la propiedad 3°.

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_n(t)\,dt = \frac{1}{n+1} \sum_{k=0}^{n} \frac{1}{\pi} \int_{-\pi}^{\pi} D_k(t)\,dt = 1.$$

Por cuanto el núcleo $\Phi_n(t)$ es par, de aquí se infiere que

$$\frac{2}{\pi} \int_{0}^{\pi} \Phi_n(t)\,dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_n(t)\,dt = 1.$$

Demostraremos ahora la propiedad 4°, de la cual se deduce, evidentemente, la propiedad 2°. Para $t \neq 2k\pi$, $k = 0, \pm 1, \pm 2, \ldots$, tenemos

$$(n+1)\Phi_n(t) = \sum_{k=0}^{n} D_k(t) = \sum_{k=0}^{n} \frac{\operatorname{sen}\left(k + \frac{1}{2}\right) t}{2\operatorname{sen}\frac{t}{2}} =$$

$$= \sum_{k=0}^{n} \frac{2\operatorname{sen}\frac{t}{2} \operatorname{sen}\left(k + \frac{1}{2}\right) t}{4\operatorname{sen}^2 \frac{t}{2}} = \frac{1}{4\operatorname{sen}^2 \frac{t}{2}} \sum_{k=0}^{n} [\cos kt - \cos(k+1)t] =$$

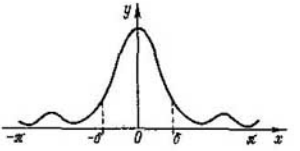


Fig. 235

$$= \frac{1 - \cos(n+1)t}{4\operatorname{sen}^2\frac{t}{2}} = \frac{\operatorname{sen}^2\frac{n+1}{2}t}{2\operatorname{sen}^2\frac{t}{2}} \quad \square$$

Observemos que la fórmula (55.37) puede obtenerse, además de la propiedad 4°, pasando al límite y haciendo tender t hacia cero y este procedimiento es factible en vista de que el núcleo de Fejer es continuo.

DEMOSTRACIÓN DEL COROLARIO. Haciendo uso de la propiedad 4° del núcleo de Fejer, obtenemos

$$0 \leqslant \max_{\delta \leqslant |t| \leqslant \pi} \Phi_n(t) = \frac{1}{2(n+1)} \max_{\delta \leqslant |t| \leqslant \pi} \frac{\operatorname{sen}^2\frac{n+1}{2}t}{\operatorname{sen}^2\frac{t}{2}} \leqslant \frac{1}{2(n+1)\operatorname{sen}^2\frac{\delta}{2}}.$$

Por cuanto el segundo miembro de la desigualdad obtenida tiende a cero para $n \to \infty$, de la estimación obtenida se deduce directamente (55.38). $\square$

La forma aproximada de la gráfica del núcleo de Fejer está expuesta en la fig. 235.

En este punto sólo se considerarán aquellas funciones f que son continuas en el segmento $[-\pi, \pi]$ y toman en sus extremos valores iguales: $f(-\pi) = f(\pi)$. Es obvio que toda función de tal índole puede ser prolongada de modo 2π-periódico desde el segmento $[-\pi, \pi]$ a todo el eje numérico R. La función que se obtendrá, la designaremos con $\bar{f}$, será continua en todo el eje R.

La función de partida f, como cualquiera función continua en un segmento, es acotada, es decir, existe una constante $M > 0$ que $|f(x)| \leqslant M$, $x \in [-\pi, \pi]$. Está claro que en este caso

$$|\bar{f}(x)| \leqslant M, \quad x \in R,$$

es decir, la función $\bar{f}$ está acotada en todo el eje R.

Además, la función $\bar{f}$ es uniformemente continua en todo el eje R. Efectivamente, siendo continua en cualquier segmento finito, por ejemplo, en $[0, 4\pi]$, es en dicho segmento uniformemente continua (véase el teorema 5 en el p. 19.7). Esto significa que para todo $\varepsilon > 0$ existe tal δ, $0 < \delta < 2\pi$ que para cualesquiera $x_1 \in [0, 4\pi]$, $x_2 \in [0, 4\pi]$, $|x_2 - x_1| < \delta$, se verifica la desigualdad

$$|\bar{f}(x_2) - \bar{f}(x_1)| < \varepsilon.$$

Pero, para x_1' y x_2' arbitrarios de tal género que $|x_2' - x_1'| < \delta$, se encontrarán unos números enteros n y m, para los cuales $x_1 \overset{\text{def}}{=} x_1' - 2\pi n \in [0, 4\pi]$, $x_2 \overset{\text{def}}{=} x_2' - 2\pi m \in [0, 4\pi]$ y $|x_1 - x_2| < \delta$, y, por cuanto $\bar{f}(x_1) = \bar{f}(x_1')$, $\bar{f}(x_2) = \bar{f}(x_2')$ son 2π-periódicas, entonces

$$|\bar{f}(x_2') - \bar{f}(x_1')| = |\bar{f}(x_2) - \bar{f}(x_1)| < \varepsilon.$$

Esto precisamente es indicio de que la función $\bar{f}$ es uniformemente continua en todo el eje numérico R.

En lo que sigue, una función prolongada de modo periódico se designará con el mismo símbolo f que la función prolongable.

Teorema 6 (de Fejer). *Si una función es continua en el segmento* $[-\pi, \pi]$ *y toma en sus extremos valores iguales, la sucesión de sus sumas de Fejer converge uniformemente en este segmento hacia la misma función.*

Corolario. *Si la serie de Fourier de una función, que es continua en el segmento* $[-\pi, \pi]$ *y toma en sus extremos valores iguales, converge en cierto punto, converge hacia el valor de la función en dicho punto.*

DEMOSTRACIÓN. Supongamos que la función f es continua en el segmento $[-\pi, \pi]$ y $f(-\pi) = f(\pi)$. Prolonguémosla de modo 2π-periódico a todo el eje numérico R. Estimemos la diferencia $f(x) - \sigma_n(x)$ entre la función f y su suma de Fejer σ_n, empleando con este fin la representación de la suma de Fejer en la forma (55.36) y las propiedades del núcleo de Fejer demostradas en el lema 6 y en su corolario. Elijamos arbitrariamente $\varepsilon > 0$. Tenemos

$$|f(x) - \sigma_n(x)| = \left| \frac{f(x)}{\pi} \int_{-\pi}^{\pi} \Phi_n(t)\,dt - \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_n(t) f(x+t)\,dt \right| =$$

$$= \frac{1}{\pi} \left| \int_{-\pi}^{\pi} \Phi_n(t)[f(x) - f(x+t)]\,dt \right| \leqslant$$

$$\leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} \Phi_n(t)\,|f(x) - f(x+t)|\,dt = \frac{1}{\pi} \int_{-\pi}^{-\delta} + \frac{1}{\pi} \int_{-\delta}^{\delta} + \frac{1}{\pi} \int_{\delta}^{\pi}, \quad (55.39)$$

donde $\delta > 0$ se ha elegido de una manera tal que el valor del módulo de continuidad $\omega(\delta; f)$ de la función f satisface la desigualdad

$$\omega(\delta; f) < \frac{\varepsilon}{3}.$$

Esto es bien posible, puesto que la función f es uniformemente continua en todo el eje numérico R. Por eso, para todo $x \in R$:

$$\frac{1}{\pi} \int_{-\delta}^{\delta} \Phi_n(t)\,|f(x) - f(x+t)|\,dt \leqslant \frac{\omega(\delta; f)}{\pi} \int_{-\delta}^{\delta} \Phi_n(t)\,dt \leqslant \frac{\varepsilon}{3\pi} \int_{-\pi}^{\pi} \Phi_n(t)\,dt = \frac{\varepsilon}{3}. \quad (55.40)$$

Las dos integrales restantes se estiman de un modo igual: la función f está acotada en toda la recta numérica, es decir, existe tal constante $M > 0$ que para cualesquiera $x \in R$ tiene lugar la desigualdad

$$|f(x)| \leqslant M.$$

Por consiguiente, para todo $x \in R$:

$$\frac{1}{\pi}\int_{\delta}^{\pi} \Phi_n(t)|f(x) - f(x+t)|\,dt \leqslant \frac{1}{\pi}\int_{\delta}^{\pi} \Phi_n(t)[|f(x)| + |f(x+t)|]\,dt \leqslant$$

$$\leqslant \frac{2M}{\pi}\int_{\delta}^{\pi} \Phi_n(t)dt \leqslant \frac{2M}{\pi}\max_{\delta \leqslant t \leqslant \pi} \Phi_n(t)\int_{\delta}^{\pi} dt =$$

$$= \frac{2M(\pi - \delta)}{\pi}\max_{\delta \leqslant t \leqslant \pi} \Phi_n(t) < 2M \max_{\delta \leqslant t \leqslant \pi} \Phi_n(t).$$

De acuerdo con el corolario del lema 6, el segundo miembro de la desigualdad obtenida tiende hacia cero cuando $n \to \infty$, por lo cual existe tal n_0 que para cualesquiera $n \geqslant n_0$ se cumple la desigualdad

$$\frac{1}{\pi}\int_{\delta}^{\pi} \Phi_n(t)|f(x) - f(x+t)|\,dt < \frac{\varepsilon}{3}. \tag{55.41}$$

Por analogía, para todo $x \in R$ y cualesquiera $n \geqslant n_0$:

$$\frac{1}{\pi}\int_{-\pi}^{-\delta} \Phi_n(t)|f(x) - f(x+t)|\,dt < \frac{\varepsilon}{3}. \tag{55.42}$$

De (55.39), (55.40), (55.41) y (55.42) tenemos para $x \in R$ arbitrario y todos los $n \geqslant n_0$:

$$|f(x) - \sigma_n(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

es decir, la sucesión $\{\sigma_n\}$ converge uniformemente en todo el eje numérico R hacia la función f. □

DEMOSTRACIÓN DEL COROLARIO. Cualquier serie convergente se suma por el método de medias aritméticas a su suma (véase el p. 35.15). Por ello, si la serie de Fourier de una función, que es continua en el segmento $[-\pi, \pi]$ y toma en sus extremos los valores iguales, converge en cierto punto hacia algún número A, entonces el límite de la sucesión de medias aritméticas de las sumas parciales, es decir, de las sumas de Fejer es también igual a A: si $\lim\limits_{n\to\infty} S_n(x_0; f) = A$, se tiene $\lim\limits_{n\to\infty} \sigma_n(x_0) = A$.

Pero, según el lema demostrado, $\lim_{n-\infty} \sigma_n(x_0) = f(x_0)$, por consiguiente, también $\lim_{n-\infty} S_n(x_0; f) = f(x_0)$. □

Subrayemos que la serie de Fourier de una función, que es continua en el segmento $[-\pi, \pi]$ y toma en sus extremos los valores iguales, puede divergir en algunos puntos. No obstante, de acuerdo con lo demostrado, si la serie converge en cierto punto, converge necesariamente hacia el valor de la misma función en dicho punto.

Observemos en conclusión que para una función, que es continua en un segmento y toma en sus extremos valores iguales, la serie de Fourier, sea convergente o divergente en los puntos aislados, permite restablecer unívocamente la función indicada: resulta suficiente formar, sirviéndose de sus sumas parciales, las sumas de Fejer; la sucesión de las últimas ya converge y, además, uniformemente hacia la propia función. De este modo, incluso el estudio de la serie divergente puede resultar útil.

55.7. APROXIMACIÓN DE LAS FUNCIONES CONTINUAS POR MEDIO DE LOS POLINOMIOS

Definición 10. *Las funciones de la forma*

$$\frac{A_0}{2} + \sum_{k=1}^{n} A_k \cos kx + B_k \operatorname{sen} kx \quad A_n^2 + B_n^2 > 0$$

se llaman polinomios trigonométricos de n-ésimo orden, n = 0, 1, 2, . . . [*].

Teorema 7 (de Weierstrass). *Si la función f es continua en el segmento* $[-\pi, \pi]$ *y* $f(-\pi) = f(\pi)$, *entonces para todo número* $\varepsilon > 0$ *existe un polinomio trigonométrico* $T(x)$ *tal que*

$$|f(x) - T(x)| < \varepsilon, \quad -\pi \leqslant x \leqslant \pi.$$

En efecto, en virtud del teorema 6 (véase el p. 55.6), a título de tal polinomio trigonométrico puede tomarse, por ejemplo, la suma correspondiente de Fejer $\sigma_n(x)$, que es, evidentemente, un polinomio trigonométrico de orden no superior a n.

Teorema 8 (de Weierstrass). *Si la función f es continua en el segmento* $[a, b]$, *para todo* $\varepsilon > 0$ *existe tal polinomio algebraico* $P(x)$ *que*

$$|f(x) - P(x)| < \varepsilon, \quad a \leqslant x \leqslant b.$$

DEMOSTRACIÓN. Apliquemos linealmente el segmento $[0, \pi]$ sobre otro segmento $[a, b]$:

$$x = a + \frac{b-a}{\pi} t, \quad 0 \leqslant t \leqslant \pi, \quad a \leqslant x \leqslant b,$$

y sea $f^*(t) = f\left(a + \dfrac{b-a}{\pi} t\right)$. La función f^* está definida por esta fórmula en $[0, \pi]$. Prolongámosla de modo impar al segmento $[-\pi, 0]$, es decir, pongamos

$$f^*(t) = f^*(-t), \quad \text{si} \quad t \in [-\pi, 0].$$

---

[*] Aquí se considera que $B_0 = 0$.

La función f^*, obtenida de esta manera, es continua en $[-\pi, \pi]$ (¿por qué?) y $f^*(-\pi) = f^*(\pi)$. Por eso, de acuerdo con el teorema 7, para cualquier número $\varepsilon > 0$ existe un polinomio trigonométrico $T(t)$ tal que

$$|f^*(t) - T(t)| < \frac{\varepsilon}{2}.$$

Como se sabe, $\cos kt$ y $\operatorname{sen} kt$, $k = 1, 2, \ldots$, y, por lo tanto, también el polinomio trigonométrico $T(t)$ son funciones analíticas, a consecuencia de lo cual se desarrollan en las series de potencias que convergen en toda la recta real y, por consiguiente, convergen uniformemente en todo segmento finito (véase el § 37):

$$T(t) = \sum_{k=0}^{\infty} c_k t^k$$

Si $P_n(t)$ son sumas parciales de esta serie, debido a su convergencia uniforme en el segmento $[-\pi, \pi]$, existe tal número n_ε que para $n \geqslant n_\varepsilon$ se tiene

$$|T(t) - P_n(t)| < \frac{\varepsilon}{2}, \quad -\pi \leqslant t \leqslant \pi.$$

Al tomar, para concretar, $n = n_\varepsilon$ y suponer $P(t) = P_{n_\varepsilon}(t)$, tenemos

$$|f^*(t) - P(t)| \leqslant |f^*(t) - T(t)| + |T(t) - P(t)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon.$$

Volviendo a la variable x, es decir, suponiendo $t = \pi \dfrac{x - a}{b - a}$ obtendremos

$$\left| f(x) - P\left(\pi \frac{x - a}{b - a}\right) \right| < \varepsilon, \quad a \leqslant x \leqslant b,$$

donde $P\left(\pi \dfrac{x - a}{b - a}\right)$ es, evidentemente, un polinomio. □

OBSERVACIÓN. Supongamos que la función f es continua en el segmento $[a, b]$. Elijamos una sucesión de los números $\varepsilon_n > 0$, $n = 1, 2, \ldots$, que tiende hacia cero $\left(\text{por ejemplo, } \varepsilon_n = \dfrac{1}{n}\right)$; entonces, de acuerdo con el teorema 8, para todo $n = 1, 2, \ldots$ existe un polinomio $P_n(x)$ (aquí n es número de orden y no el grado del polinomio), tal que

$$|f(x) - P_n(x)| < \varepsilon_n, \quad a \leqslant x \leqslant b. \tag{55.43}$$

Es evidente que para $n \to \infty$ se tiene $P_n(x) \rightrightarrows f(x)$ en el segmento $[a, b]$.

Así pues, toda función continua en un segmento sirve de límite para la sucesión de polinomios uniformemente convergente en dicho segmento. Anteriormente (véase el teorema 8' en el p. 36.4) ya se ha demostrado la afirmación recíproca, es decir, toda función que sirve de límite para una sucesión de polinomios uniformemente convergente en cierto segmento (y más aún, para una sucesión de cualesquiera funciones continuas) es continua en dicho segmento.

De este modo, el teorema de Weierstrass establece la propiedad característica de las funciones continuas, y sólo continuas.

Resulta curioso notar que el concepto original de continuidad de una función fue introducido en la forma general abstracta; este concepto de ningún modo estaba ligado con las clases concretas de funciones elementales, en particular, con los polinomios y, consecuentemente, con ningunas representaciones analíticas de las funciones en términos de los polinomios.

El teorema de Weierstrass muestra que la clase de funciones continuas introducida de este modo no está lejos, en cierto sentido, de la clase de polinomios. A saber, cualquiera que sea la función f, continua en un segmento, y por pequeño que sea el número $\varepsilon > 0$, prefijado de antemano, siempre existe un polinomio que en todo el segmento se diferencia de la función f a lo sumo en ε, es decir, un polinomio que aproxima la función con cualquier grado de exactitud dado de antemano. No es difícil obtener también la representación analítica en forma de una serie de polinomios para la función continua en un segmento. De (55.43) tenemos

$$f(x) = \lim_{n \to \infty} P_n(x), \quad a \leqslant x \leqslant b, \tag{55.44}$$

o bien

$$f(x) = P_1(x) + \sum_{n=1}^{\infty} [P_{n+1}(x) - P_n(x)] \tag{55.45}$$

($P_n(x)$ son los polinomios), con la particularidad de que en el segmento $[a, b]$ los polinomios $P_n(x)$ en (55.44) tienden a su límite uniformemente y también uniformemente converge la série (55.45) en dicho segmento. En este caso, tanto la existencia del límite (55.44), como la del desarrollo (55.45) constituyen una condición necesaria y suficiente de la continuidad de la función f en el segmento citado. Esto justifica la idea intuitiva sobre una función como expresión analítica compuesta de una variable independiente y unas constantes por medio de las operaciones algebraicas y analíticas.

Observaciones análogas pueden hacerse también sobre el primer teorema de Weierstrass (teorema 7).

55.8. COMPLETITUD DEL SISTEMA TRIGONOMÉTRICO Y DEL SISTEMA DE POTENCIAS ENTERAS NO NEGATIVAS DE X EN UN ESPACIO DE FUNCIONES CONTINUAS

En este punto se parafrasearán los teoremas demostrados anteriormente y se deducirán de éstos algunos corolarios sencillos.

Definición 11. *Sea X un conjunto de las funciones definidas en el segmento $[a, b]$. Un sistema de funciones*

$$\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots \tag{55.46}$$

se llama completo para el conjunto X en el sentido de aproximación uniforme, si, cualquiera que sea la función $f \in X$, para todo $\varepsilon > 0$ existe tal número finito de funciones $\varphi_{n_1}, \varphi_{n_2}, \ldots, \varphi_{n_k}$, pertenecientes al sistema (55.46), y tales números $\lambda_1, \lambda_2, \ldots, \lambda_k$ que

$$|f(x) - [\lambda_1 \varphi_{n_1}(x) + \lambda_2 \varphi_{n_2}(x) + \ldots + \lambda_k \varphi_{n_k}(x)]| < \varepsilon$$

para cualesquiera $x \in [a, b]$.

En otras palabras, el sistema de funciones (55.46) forma un sistema completo para el conjunto X, si toda función de X puede ser aproximada con la exactitud requerida mediante las combinaciones lineales de funciones del sistema (55.46).

Haciendo uso del concepto de completitud de un sistema, los teoremas 7 y 8 del párrafo precedente pueden parafrasearse del modo siguiente.

Teorema 7'. *El sistema de funciones trigonométricas* (55.2) *es completo, en el sentido de aproximación uniforme, para el conjunto de funciones que son continuas en el segmento* $[-\pi, \pi]$ *y toman en sus extremos valores iguales.*

Teorema 8'. *Un sistema de las potencias enteras no negativas de* x, *es decir, el sistema*

$$1, x, x^2, \ldots, x^n, \ldots \tag{55.47}$$

es completo, en el sentido de aproximación uniforme, para el conjunto de todas las funciones que son continuas en cualquier segmento dado.

Definición 12. *Supongamos que las funciones* f *y* g *vienen definidas en el segmento* $[a, b]$. *El número*

$$\sqrt{\int_a^b [f(x) - g(x)]^2 dx}$$

se denomina desviación estándar en el segmento $[a, b]$ *de la función* f *de la* g. *)

Definición 13. *El sistema de funciones* (55.46) *se llama completo, en el sentido de la aproximación media cuadrática, para cierto conjunto* X *de funciones definidas en el segmento* $[a, b]$, *si, cualquiera que sea la función* $f \in X$, *para todo* $\varepsilon > 0$ *existe tal combinación lineal finita de funciones del sistema* (55.46) *que su desviación estándar en el segmento* $[a, b]$ *de la función* f *es inferior a* ε.

Teorema 9. *El sistema de funciones trigonométricas* (55.2) *es completo en el sentido de la aproximación media cuadrática en un conjunto de las funciones que son continuas en el segmento* $[-\pi, \pi]$ *y toman en los puntos* π *y* $-\pi$ *un mismo valor.*

DEMOSTRACIÓN. Sea f una función continua en el segmento $[-\pi, \pi]$, con la particularidad de que $f(\pi) = f(-\pi)$. De conformidad con el teorema 7', para todo $\varepsilon > 0$ existe un polinomio trigonométrico $T(x)$ tal que

$$|f(x) - T(x)| < \frac{\varepsilon}{\sqrt{2\pi}}, \quad -\pi \leqslant x \leqslant \pi.$$

De aquí para la desviación estándar de este polinomio de la función f tenemos

$$\sqrt{\int_{-\pi}^{\pi} [f(x) - T(x)]^2 dx} < \frac{\varepsilon}{\sqrt{2\pi}} \sqrt{\int_{-\pi}^{\pi} dx} = \varepsilon. \ \square$$

En lo que sigue veremos (véase el p. 58.6) que la acotación $f(\pi) = f(-\pi)$, utilizada por nosotros en la demostración del teorema 9 (sólo en este caso se podría referir al teorema 7'), no es esencial. A saber, el sistema trigonométrico (55.2) es completo en el sentido de la media cuadrática en todo el conjunto de funciones con-

*) Podemos decir también "desviación de la función g de la f", por cuanto la expresión en consideración no cambia su valor, si f y g cambian de lugar.

tinuas en el segmento $[-\pi, \pi]$, y, más aún, se puede mostrar que es completo también en el sentido de la media cuadrática en el conjunto de todas las funciones de cuadrado integrable en el segmento $[-\pi, \pi]$.

Observemos que el sistema trigonométrico (55.2) es incompleto a ciencia cierta en el conjunto de todas las funciones continuas en el segmento $[-\pi, \pi]$ en el sentido de la aproximación uniforme, es decir, en el sentido de la definición 11. En efecto, si f es una función tal que para todo $\varepsilon > 0$ existe un polinomio trigonométrico T_ε de tal índole que

$$|f(x) - T_\varepsilon(x)| < \varepsilon, \quad -\pi \leqslant x \leqslant \pi,$$

entonces de la condición $T_\varepsilon(\pi) = T_\varepsilon(-\pi)$ para $\varepsilon \to 0$ proviene que $f(\pi) = f(-\pi)$.

En la aproximación de las funciones en el sentido de la media cuadrática mediante los polinomios trigonométricos desempeñan un papel especial las sumas parciales de la serie de Fourier de la función que se aproxima. En el punto que sigue se mostrará que la suma parcial de n-ésimo orden tiene desviación estándar mínima de la función dada en comparación con cualquier polinomio trigonométrico de grado n.

En fin, se puede mostrar que si la función f posee el cuadrado integrable en el segmento $[-\pi, \pi]$, entonces la desviación, que experimentan sus sumas parciales de Fourier $S_n(x)$ tiende a cero cuando $n \to \infty$, o, como suele decirse, la función f de cuadrado integrable es un límite, en el sentido de la media cuadrática, de sus sumas parciales de Fourier (véase el p. 58.6). Todas estas circunstancias atestiguan en favor del estudio de la aproximación de funciones en el sentido de la desviación estándar.

Por analogía con el teorema 9 se demuestra el teorema siguiente.

Teorema 10. *El sistema de potencias enteras y no negativas de x, es decir, el sistema* (55.47) *es completo en el sentido de la aproximación media cuadrática en un conjunto de funciones continuas sobre cualquier segmento dado.*

DEMOSTRACIÓN. Sea f una función continua en cierto segmento $[a, b]$. Entonces, para todo $\varepsilon > 0$, de acuerdo con el teorema 8′, existe tal polinomio P que

$$|f(x) - P(x)| < \frac{\varepsilon}{\sqrt{b-a}}, \quad a \leqslant x \leqslant b,$$

de donde

$$\sqrt{\int_a^b [f(x) - P(x)]^2 dx} < \varepsilon. \ \square$$

55.9. PROPIEDAD MINIMAL DE LOS COEFICIENTES DE FOURIER. DESIGUALDAD DE BESSEL E IGUALDAD DE PARSEVAL

En este punto examinaremos las series de Fourier para unas funciones integrables cuyo cuadrado es también integrable en el segmento $[-\pi, \pi]$ (la integrabilidad aquí se entiende, por regla general, en el sentido impropio). Es importante indicar que si la función f es de tal género que tiene un número finito de puntos singu-

lares (véase el p. 55.1) en el segmento $[-\pi, \pi]$, es integrable según Riemann en cualquier segmento que no contiene ningún punto singular y el cuadrado de ésta f^2 es integrable en el segmento $[-\pi, \pi]$, entonces de la desigualdad

$$|f| \leqslant \frac{1 + |f|^2}{2}$$

se infiere que la función f es integrable en el segmento citado. Lo recíproco, en el caso general, no es cierto. Existen funciones positivas $\left(\text{por ejemplo, la función } \frac{1}{\sqrt{|x|}}\right)$, integrables en el segmento $[-\pi, \pi]$ cuyo cuadrado, no obstante, ya no es integrable en el mismo.

De este modo, el conjunto citado de funciones de cuadrado integrable en el segmento $[-\pi, \pi]$ constituye un subconjunto propio del conjunto de todas las funciones absolutamente integrables en el segmento $[-\pi, \pi]$.

Observemos que análogamente se introduce también el concepto de función de cuadrado integrable para cualquier intervalo finito.

Teorema 11. *Sea f una función de cuadrado integrable en el segmento $[-\pi, \pi]$. Entonces, si $S_n(x)$ es su suma de Fourier de n-ésimo orden, se tiene*

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = \min_{T_n(x)} \int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 dx, \tag{55.48}$$

donde el mínimo en el segundo miembro de la igualdad se toma según todos los polinomios trigonométricos T_n de orden no superior a n.

Si $a_0, a_n, b_n, n = 1, 2, \ldots$, son coeficientes de Fourier de la función f, se verifica la desigualdad

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2 \leqslant \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx, \tag{55.49}$$

que lleva el nombre de Bessel[*].

DEMOSTRACIÓN. Sea

$$T_n(x) = \frac{A_0}{2} \sum_{k=1}^{n} A_k \cos kx + B_k \operatorname{sen} kx,$$

en este caso, abriendo los corchetes en la expresión

$$\int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 dx \tag{55.50}$$

y haciendo uso del lema 1 del p. 55.1 (en particular, la ortogonalidad del sistema trigonométrico), obtenemos

$$\int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 dx = \int_{-\pi}^{\pi} f^2(x) dx + \pi \left(\frac{A_0^2}{2} + \sum_{k=1}^{n} A_k^2 + B_k^2 \right) -$$

[*] F. Bessel (1784 — 1846), matemático y astrónomo alemán.

$$- 2\left[\frac{A_0}{2}\int_{-\pi}^{\pi} f(x)dx + \sum_{k=1}^{n} A_k \int_{-\pi}^{\pi} f(x)\cos kx\,dx + \right.$$

$$\left. + B_k \int_{-\pi}^{\pi} f(x)\operatorname{sen} kx\,dx\right] = \int_{-\pi}^{\pi} f^2(x)dx + \pi\left(\frac{A_0^2}{2} + \sum_{k=1}^{n} A_k^2 + B_k^2\right) -$$

$$- 2\pi\left[\frac{a_0 A_0}{2} + \sum_{k=1}^{n} a_k A_k - b_k B_k\right] = \int_{-\pi}^{\pi} f^2(x)dx +$$

$$+ \pi\left[\frac{(A_0 - a_0)^2}{2} + \sum_{k=1}^{n} (A_k - a_k)^2 + (B_k - b_k)^2\right] -$$

$$- \pi\left[\frac{a_0}{2} + \sum_{k=1}^{n} a_k^2 + b_k^2\right]. \quad (55.51)$$

De la expresión obtenida se ve que la magnitud (55.50) adquiere el valor mínimo cuando $A_0 = a_0$, $A_k = a_k$, $B_k = b_k$, $k = 1, 2, \ldots$, es decir, cuando $T_n(x)$ es la suma de Fourier $S_n(x)$ de orden n de la función f. La primera afirmación del teorema queda demostrada.

Si $T_n(x)$ es la suma de Fourier de orden n, de (55.51) se desprende que

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = \int_{-\pi}^{\pi} f^2(x)dx - \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{n} a_k^2 + b_k^2\right], \quad (55.52)$$

de donde

$$\int_{-\pi}^{\pi} f^2(x)dx - \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{n} a_k^2 + b_k^2\right] \geqslant 0.$$

Esta desigualdad es lícita para cualquier n natural. Pasando en la misma al límite para $n \to \infty$, obtendremos una desigualdad

$$\int_{-\pi}^{\pi} f^2(x)dx - \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{\infty} a_k^2 + b_k^2\right] \geqslant 0,$$

la que, obviamente, es equivalente a la desigualdad (55.49). □

De la desigualdad de Bessel se deduce que para una función de cuadrado integrable la serie

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2$$

converge. El término general de la serie convergente tiende a cero, por lo cual en el caso dado $\lim_{n-\infty} a_n = \lim_{n-\infty} b_n = 0$.

De este modo hemos establecido nuevamente que los coeficientes de Fourier tienden a cero (véase el p. 55.2), pero esta vez para una clase de funciones más estrecha (como lo indicamos al principio de este punto) que antes, a saber, para la clase de funciones de cuadrado integrable.

En el punto 58.6 se mostrará que en realidad la fórmula (55.49) queda lícita con *el signo de igualdad*. Por ahora demostraremos este hecho sólo para el caso en que la función f es continua y 2π-periódica.

Teorema 12. *Supongamos que la función f es continua en el segmento $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ y $a_0, a_n, b_n, n = 1, 2, \ldots$, son sus coeficientes de Fourier. En este caso se verifica la igualdad*

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2,$$

*llamada igualdad de Parseval**).

DEMOSTRACIÓN. Debido a la completitud en el sentido de aproximación media cuadrática del sistema de funciones trigonométricas (55.2) en la clase de funciones continuas que toman valores iguales en los extremos del segmento $[-\pi, \pi]$, siendo todo $\varepsilon > 0$, existe para la función f un polinomio trigonométrico $T(x)$ de cierto orden k tal que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} [f(x) - T(x)]^2 dx < \varepsilon. \tag{55.53}$$

Con arreglo al teorema 11 (véase (55.48)), para la suma de Fourier $S_k(x)$ del mismo orden k se cumple la desigualdad

$$\int_{-\pi}^{\pi} [f(x) - S_k(x)]^2 dx \leqslant \int_{-\pi}^{\pi} [f(x) - T(x)]^2 dx.$$

De aquí y de las fórmulas (55.52) y (55.53) obtenemos:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)dx - \left[\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2\right] \leqslant$$

$$\leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)dx - \left[\frac{a_0^2}{2} + \sum_{n=1}^{k} a_n^2 + b_n^2\right] =$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} [f(x) - S_k(x)]^2 dx \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} [f(x) - T(x)]^2 < \varepsilon.$$

---

*) M. Parseval (1755 — 1836), matemático francés.

Por cuanto esta desigualdad es válida para cualquier $\varepsilon > 0$, su primer miembro es igual a cero. □

Corolario. *Cumplidas las suposiciones del teorema, tenemos*

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = 0.$$

En efecto, en virtud del teorema 12, para $n \to \infty$ el segundo miembro de la igualdad (55.52) tiende a cero. □

55.10. CARÁCTER DE CONVERGENCIA DE LAS SERIES DE FOURIER. DERIVACIÓN DE LAS SERIES DE FOURIER TÉRMINO A TÉRMINO

Analicemos la relación que existe entre las series de Fourier de una función y la derivada de ésta.

Teorema 13. *Sea f una función continua en el segmento $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ y supongamos que*

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

Si la función f es continuamente derivable a trozos en el segmento $[-\pi, \pi]$ (véase la definición 1 en el p. 30.2), entonces

$$f'(x) \sim \sum_{n=1}^{\infty} -n a_n \operatorname{sen} nx + n b_n \cos nx,$$

es decir, la serie de Fourier de la derivada se obtiene de la serie de Fourier de la misma función por derivación formal término a término *).

DEMOSTRACIÓN. Sea

$$f'(x) \sim \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} \alpha_n \cos nx + \beta_n \operatorname{sen} nx.$$

Teniendo presente que $f(\pi) = f(-\pi)$ e integrando por partes, obtenemos

$$\alpha_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(t)\,dt = \frac{1}{\pi} [f(\pi) - f(-\pi)] = 0,$$

$$\alpha_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(t) \cos nt\,dt = \frac{1}{\pi} f(t) \cos nt \Big|_{-\pi}^{\pi} + \frac{n}{\pi} \int_{-\pi}^{\pi} f(t) \operatorname{sen} nt\,dt = n b_n,$$

$$\beta_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(t) \operatorname{sen} nt\,dt = f(t) \operatorname{sen} nt \Big|_{-\pi}^{\pi} -$$

*) Sin suposiciones algunas acerca de la convergencia de la serie de Fourier de la derivada.

$$-\frac{n}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt = -na_n, \quad n = 1, 2, \ldots. \quad \square$$

Pasemos al estudio de la velocidad de convergencia de la serie de Fourier en dependencia de la suavidad de las funciones. Previamente demostremos el lema.

Lema 7. *Supongamos que la función f tiene en un segmento derivadas continuas de orden hasta $k-1$ inclusive y una derivada continua a trozos de orden k ($k \geqslant 1$)*[*], *con la particularidad de que*

$$f^{(j)}(-\pi) = f^{(j)}(\pi), \quad j = 0, 1, \ldots, k-1,$$

y sea

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos nx + b_n\operatorname{sen} nx.$$

En este caso

$$|a_n| \leqslant \frac{\varepsilon_n}{n^k}, \quad |b_n| \leqslant \frac{\varepsilon_n}{n^k}, \quad n = 1, 2, \ldots,$$

donde $\varepsilon_n > 0$ y la serie $\sum_{n=1}^{\infty} \varepsilon_n^2$ converge.

DEMOSTRACIÓN. Al aplicar sucesivamente el teorema 13 k veces, obtendremos

$$f^{(k)}(x) \sim \sum_{n=1}^{\infty} \alpha_n\cos nx + \beta_n\operatorname{sen} nx,$$

donde o bien

$$\alpha_n = \pm n^k a_n, \quad \beta_n = \pm n^k b_n, \tag{55.54}$$

o bien

$$\alpha_n = \pm n^k b_n, \quad \beta_n = \pm n^k a_n, \tag{55.55}$$

con la particularidad de que, según la desigualdad de Bessel,

$$\sum_{n=1}^{\infty} \alpha_n^2 + \beta_n^2 \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} [f^{(k)}(x)]^2 dx. \tag{55.56}$$

Pongamos $\varepsilon_n = \sqrt{\alpha_n^2 + \beta_n^2}$. En virtud de la desigualdad (55.56), la serie $\sum_{n=1}^{\infty} \varepsilon_n^2$ converge.

---

[*] Decimos que cierta función tiene derivada continua a trozos en el segmento dado, si dicha función es continuamente derivable a trozos en el mismo (véase la definición 1 en el p. 30.2). De este modo, si la función tiene derivada continua a trozos en un segmento, puede suceder que en un número finito de puntos del mismo segmento ella no tenga en absoluto derivada. Por ejemplo, la función $f(x) = |x|$ en el segmento $[-1, 1]$ tiene derivada continua a trozos, mientras que en el punto $x = 0$ no la tiene.

Si es lícita (55.54), se tiene

$$|a_n| = \frac{|\alpha_n|}{n^k} \leqslant \frac{\sqrt{\alpha_n^2 + \beta_n^2}}{n^k} = \frac{\varepsilon_n}{n^k}.$$

Análogamente,

$$|b_n| \leqslant \frac{\varepsilon_n}{n^k}, \quad k = 1, 2, \ldots .$$

De manera semejante esta estimación se obtiene para el caso (55.55). □

Teorema 14. *Supongamos que la función f tiene en el segmento $[-\pi, \pi]$ derivadas continuas hasta el orden $k-1$ inclusive y una derivada continua a trozos de orden $k (k \geqslant 1)$, con la particularidad de que $f^{(j)}(-\pi) = f^{(j)}(\pi)$, $j = 0, 1, \ldots, k-1$. Entonces, la serie de Fourier de la función f converge uniforme y absolutamente en todo el segmento $[-\pi, \pi]$ a la propia función f y*

$$|f(x) - S_n(x; f)| \leqslant \frac{\eta_n}{n^{k-1/2}},$$

donde $\lim_{n\to\infty} \eta_n = 0$ ($\{\eta_n\}$ es una sucesión numérica), y $S_n(x; f)$ es la suma de Fourier de n-ésimo orden de la función f.

De este modo podemos decir que en el segmento $[-\pi, \pi]$ se cumple uniformemente la estimación

$$f(x) - S_n(x; f) = o\left(n^{-k+\frac{1}{2}}\right), \quad n \to \infty.$$

Observemos preliminarmente que si $\{u_n\}$ y $\{v_n\}$ son las sucesiones de números no negativos de tal índole que

$$\sum_{n=1} u_n^2 < +\infty \quad \text{y} \quad \sum_{n=1}^{\infty} v_n^2 < +\infty,$$

entonces

$$\sum_{n=1}^{\infty} u_n v_n \leqslant \sqrt{\sum_{n=1}^{\infty} u_n^2} \sqrt{\sum_{n=1}^{\infty} v_n^2}. \tag{55.57}$$

En efecto, esta igualdad se obtiene inmediatamente, pasando al límite, de la desigualdad de Cauchy — Schwartz $\sum_{n=1}^{N} u_n v_n \leqslant \sqrt{\sum_{n=1}^{N} u_n^2} \sqrt{\sum_{n=1}^{N} v_n^2}$ cuando $N \to \infty$ (véase el p. 18.1 y 35.8*) (indiquemos que la desigualdad (55.57) es un caso particular de la desigualdad (35.33) del p. 35.8* para $p = q = 2$).

DEMOSTRACIÓN DEL TEOREMA 14. Sea

$$f(x) \sim \frac{a}{2} + \sum_{m=1}^{\infty} a_m \cos mx + b_m \operatorname{sen} mx, \tag{55.58}$$

$$S_n(x; f) = \frac{a_0}{2} + \sum_{m=1}^{n} a_m \cos mx + b_m \operatorname{sen} mx.$$

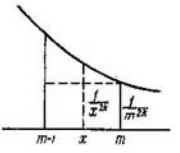


Fig. 236

Según el lema,

$$|a_m| \leqslant \frac{\varepsilon_m}{m^k}, \quad |b_m| \leqslant \frac{\varepsilon_m}{m^k}, \tag{55.59}$$

donde ε_m son de tal género que la serie

$$\sum_{m=1}^{\infty} \varepsilon_m^2 \tag{55.60}$$

converge.

Aplicando las desigualdades (55.57) y (55.59), estimemos el resto $r_n(x)$ de la serie (55.58):

$$|r_n(x)| = \left| \sum_{m=n+1}^{\infty} a_m \cos mx + b_m \operatorname{sen} mx \right| \leqslant$$

$$\leqslant \sum_{m=n+1}^{\infty} |a_m| + |b_m| \leqslant 2 \sum_{m=n+1}^{\infty} \frac{\varepsilon_m}{m^k} \leqslant$$

$$\leqslant 2 \sqrt{\sum_{m=n+1}^{\infty} \varepsilon_m^2} \sqrt{\sum_{m=n+1}^{\infty} \frac{2}{m^{2k}}}. \tag{55.61}$$

Pongamos

$$\varkappa_n = \sum_{m=n+1}^{\infty} \varepsilon_m^2.$$

Ya que la serie (55.60) es convergente, tenemos

$$\lim_{n \to \infty} \varkappa_n = 0. \tag{55.62}$$

Luego, indiquemos que en el segmento $[m-1, m]$ se verifica la desigualdad $\frac{1}{m^{2k}} \leqslant \frac{1}{x^{2k}}$ (fig. 236) y, por consiguiente, $\frac{1}{m^{2k}} \leqslant \int_{m-1}^{m} \frac{dx}{x^{2k}}$. Por esto

$$\sum_{m=n+1}^{\infty} \frac{1}{m^{2k}} \leqslant \sum_{m=n+1}^{\infty} \int_{m-1}^{m} \frac{dx}{x^{2k}} \leqslant \int_{n}^{+\infty} \frac{dx}{x^{2k}} = \frac{1}{(2k-1)n^{2k-1}}.$$

De este modo, de (55.61) se desprende la estimación

$$|r_n(x)| \leqslant 2\sqrt{\frac{\varkappa_n}{2k-1}}\,\frac{1}{\sqrt{n^{2k-1}}}. \tag{55.63}$$

Pongamos, por fin, $\eta_n = \dfrac{2}{\sqrt{2k-1}}\sqrt{\varkappa_n}$; en virtud de (55.62) $\lim\limits_{n\to\infty}\eta_n = 0$. Por ello, de la igualdad (55.63) obtenemos

$$|r_n(x)| \leqslant \frac{\eta_n}{n^{k-\frac{1}{2}}} = o\left(\frac{1}{n^{k-\frac{1}{2}}}\right), \quad n = 1, 2, \ldots,$$

y aquí el infinitésimo η_n no depende del punto x.

De conformidad con el corolario 4 del teorema 4 (el p. 55.4), la serie (55.58) converge hacia la función $f(x)$, por consiguiente, $r_n(x) = f(x) - S_n(x, f)$ y, de este modo la convergencia uniforme de la serie de Fourier con la estimación mencionada queda demostrada.

Su convergencia absoluta también se demuestra, puesto que hemos obtenido la estimación (véase (55.61))

$$\sum_{m=n+1}^{\infty} |a_m| + |b_m| \leqslant \frac{\eta_n}{n^{k-\frac{1}{2}}},$$

de la que se deduce que la serie de Fourier de la función f no sólo es absolutamente convergente, sino que una serie formada de las magnitudes absolutas de sus términos y, más aún, una serie

$$\sum_{m=1}^{\infty} |a_m| + |b_m|$$

converge con la misma "velocidad" $\dfrac{\eta_n}{n^{k-\frac{1}{2}}}$. □

El teorema 14 enseña que cuanto más suave es la función f (es decir, cuanto mayor es el número de derivadas que tiene f), tanto mayor es la velocidad de convergencia hacia ella de la serie de Fourier. En este caso la desigualdad (53.63) hace posible estimar el error que se obtiene al sustituir la serie de Fourier por su n-ésima suma parcial.

De este teorema proviene, en particular, para $k = 1$, que la serie de Fourier de toda función periódica de período 2π, continua y continuamente derivable a trozos (véase el p. 30.2), converge uniformemente sobre todo el período hacia la misma función.

Ejercicios. 11. ¿Será uniformemente convergente la serie de Fourier de la función $f(x) = |x|$, $-\pi \leqslant x \leqslant \pi$? ¿Convergirá uniformemente una serie obtenida por derivación término a término de la serie de Fourier de esta función?

12. Muéstrese que la serie de Fourier de una función continua periódica lineal a trozos (la definición de la función lineal a trozos véase en el ejercicio 6 del p. 19.7) converge hacia esta función uniformemente.

13. Sirviéndose del resultado del ejercicio antecedente y del resultado del ejercicio 6, p. 19.7, demuéstrese el teorema 7 del p. 55.7 sobre la aproximación uniforme de funciones periódicas continuas mediante los polinomios trigonométricos.

55.11. INTEGRACIÓN DE LAS SERIES DE FOURIER TÉRMINO A TÉRMINO

En este punto se probará que las series de Fourier pueden integrarse término a término.

Teorema 15. *Sea f una función continua en el segmento $[-\pi, \pi]$ y supongamos que*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx \tag{55.64}$$

es su serie de Fourier. Tenemos en este caso

$$\int_0^t f(x)\,dx = \int_0^t \frac{a_0\,dx}{2} + \sum_{n=1}^{\infty} \int_0^t (a_n \cos nx + b_n \operatorname{sen} nx)\,dx =$$

$$= \frac{a_0 t}{2} + \sum_{n=1}^{\infty} \frac{a_n}{n} \operatorname{sen} nt + \frac{b_n}{n}(1 - \cos nt), \tag{55.65}$$

y la serie que figura a la derecha converge uniformemente.

Diremos que la afirmación sobre la convergencia (incluso de la convergencia uniforme) de la serie (55.65) tiene lugar sin suposiciones algunas sobre la convergencia de la serie de partida (55.64).

DEMOSTRACIÓN. Examinemos una función

$$F(t) = \int_0^t \left[f(x) - \frac{a_0}{2} \right] dx. \tag{55.66}$$

Esta función es continua en el segmento $[-\pi, \pi]$ y tiene en éste una derivada continua $F'(t) = f(t) - \dfrac{a_0}{2}$ y

$$F(\pi) - F(-\pi) = \int_{-\pi}^{\pi} f(x)\,dx - \pi a_0 = 0.$$

Por eso, en virtud del teorema 14, su serie de Fourier converge hacia ella y, además, uniformemente. Designemos sus coeficientes de Fourier mediante A_0, A_n, B_n, $n = 1, 2, \ldots$. Entonces, en vista de lo dicho

$$F(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos nt + B_n \operatorname{sen} nt. \tag{55.67}$$

Hallemos los coeficientes de esta serie. Integrando por partes, obtendremos

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\cos nt\,dt = \frac{1}{\pi}F(t)\frac{\operatorname{sen} nt}{n}\Big|_{-\pi}^{\pi} -$$

$$-\frac{1}{n\pi}\int_{-\pi}^{\pi}\left[f(t)-\frac{a_0}{2}\right]\operatorname{sen} nt\,dt = -\frac{b_n}{n}, \quad n = 1, 2, \ldots$$

Por analogía,

$$B_n = \frac{a_n}{n}, \quad n = 1, 2, \ldots.$$

Para hallar A_0, pongamos en (55.67) $t = 0$. Al notar que $F(0) = 0$, obtendremos

$$\frac{A_0}{2} + \sum_{n=1}^{\infty} A_n = 0, \quad \text{de donde} \quad \frac{A_0}{2} = \sum_{n=1}^{\infty}\frac{b_n}{n}.$$

Así pues,

$$F(t) = \sum_{n=1}^{\infty}\frac{a_n}{n}\operatorname{sen} nt + \frac{b_n}{n}(1-\cos nt).$$

De aquí y también de (55.66) proviene precisamente la fórmula (55.65) y la convergencia uniforme de la serie (55.65) es el resultado de la convergencia uniforme de la serie (55.67). □

Problema 36. Demuéstrese que la serie trigonométrica convergente $\sum_{n=2}^{\infty}\frac{\operatorname{sen} nx}{\ln n}$ no es una serie de Fourier de ninguna función absolutamente integrable.

Indiquemos que si $\int_{-\pi}^{\pi} f(x)dx = 0$ y, por ende, $a_0 = 0$, entonces como resultado de la integración término a término de la serie de Fourier de la función f se obtiene de nuevo una serie de Fourier de cierta función F, a saber, como se deduce de lo demostrado,

$$F(x) = \int_0^x f(t)dt.$$

Dado que para cualquier función primitiva Φ de la función f, continua en el segmento $[-\pi, \pi]$, es válida la fórmula de Newton — Leibniz

$$\Phi(\pi) - \Phi(-\pi) = \int_{-\pi}^{\pi} f(x)dx,$$

la condición $\int_{-\pi}^{\pi} f(x)dx = 0$ será equivalente a que todas las funciones primitivas de f toman en los extremos del segmento $[-\pi, \pi]$ valores iguales.

55.12 SERIES DE FOURIER PARA EL CASO DE UN INTERVALO ARBITRARIO. NOTACIÓN COMPLEJA DE LAS SERIES DE FOURIER

La teoría de las series trigonométricas de Fourier de las funciones 2π-periódicas se extiende fácilmente al caso de funciones periódicas de cualquier periodo $2l$. Para ello resulta suficiente aplicar el segmento $[-l, l]$ sobre el segmento $[-\pi, \pi]$ mediante la aplicación lineal:

$$y = \frac{\pi}{l}x, \quad -l \leqslant x \leqslant l, \quad -\pi \leqslant y \leqslant \pi,$$

y el problema se reducirá al caso ya considerado. Se denomina serie de Fourier de la función f de período $2l$ respecto de la variable de partida x una serie de la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l} + b_n \operatorname{sen}\frac{n\pi x}{l},$$

donde

$$a_0 = \frac{1}{l}\int_{-l}^{l} f(t)dt, \quad a_n = \frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{n\pi t}{l}\,dt,$$

$$b_n = \frac{1}{l}\int_{-l}^{l} f(t)\operatorname{sen}\frac{n\pi t}{l}\,dt, \quad n = 1, 2, \ldots .$$

En particular, si la función f es par, se tiene

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l},$$

donde

$$a_n = \frac{2}{l}\int_{0}^{l} f(t)\cos\frac{n\pi t}{l}\,dt, \quad n = 0, 1, 2, \ldots ,$$

y si f es impar, entonces

$$f(x) \sim \sum_{n=1}^{\infty} b_n \operatorname{sen}\frac{n\pi x}{l},$$

donde

$$b_n = \frac{2}{l}\int_0^l f(t)\operatorname{sen}\frac{n\pi t}{l}\,dt, \quad n = 1, 2, \ldots .$$

Como conclusión, indiquemos, además, la así llamada *notación compleja* de las series de Fourier que es de uso frecuente en las matemáticas y sus aplicaciones. Sea

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx. \tag{55.68}$$

Como se sabe (véase el p. 37.6),

$$\cos nx = \frac{1}{2}(e^{nxi} + e^{-nxi}), \tag{55.69}$$

$$\operatorname{sen} nx = \frac{1}{2i}(e^{nxi} - e^{-nxi}) = \frac{i}{2}(e^{-nxi} - e^{nxi}). \tag{55.70}$$

Al sustituir (55.69) y (55.70) en (55.68), obtendremos

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \frac{1}{2}(a_n - b_n i)e^{nxi} + \frac{1}{2}(a_n + b_n i)e^{-nxi}.$$

Suponiendo

$$c_0 = \frac{a_0}{2}, \quad c_n = \frac{1}{2}(a_n - b_n i), \quad c_{-n} = \frac{1}{2}(a_n + b_n i),$$

tenemos

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{55.71}$$

donde, evidentemente, $c_{-n} = \bar{c}_n$, $n = 1, 2, \ldots$. Al recordar que $\cos\alpha \pm i\operatorname{sen}\alpha = e^{\pm i\alpha}$ (véase el p. 37.6), tendremos

$$c_n = \frac{1}{2}(a_n - b_n i) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(\cos nx - i\operatorname{sen} nx)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx,^{*)}$$

$$c_{-n} = \frac{1}{2}(a_n + b_n i) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}dx, \quad n = 1, 2, \ldots ,$$

o bien, al reunir ambas fórmulas y añadir el caso de $n = 0$,

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx, \quad n = 0, \pm 1, \pm 2, \ldots \tag{55.72}$$

*) Véase en el p. 54.6 la definición de la integral de una función de valores complejos.

Al sustituir (55.72) en (55.71), obtendremos

$$f(x) \sim \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} e^{inx} \int_{-\pi}^{\pi} f(t) e^{-int} dt. \tag{55.73}$$

Así pues, hemos escrito la serie de Fourier en la forma compleja y se han obtenido las expresiones para sus coeficientes.

Sólo nos resta aclarar el concepto de convergencia de la serie de forma (55.73).

Se denomina suma parcial de n-ésimo orden de la serie

$$\sum_{n=-\infty}^{+\infty} z_n \tag{55.74}$$

la suma $S_n = \sum_{k=-n}^{n} z_k$. La serie (55.74) se llama convergente, si existe $S = \lim_{n \to \infty} S_n$, en este caso S se denomina suma de la serie y se escribe

$$S = \sum_{n=-\infty}^{+\infty} z_n.$$

§ 56. INTEGRAL DE FOURIER Y TRANSFORMACIÓN DE FOURIER

56.1. REPRESENTACIÓN DE UNA FUNCIÓN EN FORMA DE LA INTEGRAL DE FOURIER

Supongamos que la función f es absolutamente integrable en todo el eje real. Escribamos para ella una integral, correspondiente, en cierto sentido, a la serie de Fourier en la que la sumación según el índice n se ha sustituido por la integración respecto de cierto parámetro:

$$\int_0^{+\infty} [a(y)\cos xy + b(y)\operatorname{sen} xy]\,dy, \tag{56.1}$$

donde

$$a(y) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t)\cos yt\,dt, \tag{56.2}$$

$$b(y) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t)\operatorname{sen} yt\,dt. \tag{56.3}$$

Las fórmulas (56.2) y (56.3) recuerdan las fórmulas para los coeficientes de Fourier.

Definición 1. *La integral* (56.1) *se llama integral de Fourier de la función* f.

Substituyendo (56.2) y (56.3) en la integral (56.1), transformemos esta última del

modo siguiente:

$$\int_0^{+\infty} [a(y)\cos xy + b(y)\operatorname{sen} xy]dy =$$

$$= \frac{1}{\pi} \int_0^{+\infty} dy \int_{-\infty}^{+\infty} f(t)(\cos ty \cos xy + \operatorname{sen} ty \operatorname{sen} xy)dt =$$

$$= \frac{1}{\pi} \int_0^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\cos y(x-t)dt. \qquad (56.4)$$

Así como la suma de una serie de Fourier de una función es igual, en ciertas condiciones, a la misma función, la integral de Fourier representa también la función de partida.

Teorema 1. *Supongamos que*

1) *la función f es continua a trozos en cada segmento finito y absolutamente integrable en toda la recta real*;

2) *en el punto x existe una derivada a la derecha* $f'_+(x)$ *y una derivada a la izquierda* $f'_-(x)$. *Entonces, para el punto citado x se verifica la fórmula*

$$\frac{f(x+0)+f(x-0)}{2} = \frac{1}{\pi} \int_0^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\cos y(x-t)dt. \qquad (56.5)$$

La fórmula (56.5) lleva el nombre de Fourier.

DEMOSTRACIÓN. Consideremos una integral

$$S(\eta) = \frac{1}{\pi} \int_0^{\eta} dy \int_{-\infty}^{+\infty} f(t)\cos y(x-t)dt, \qquad (56.6)$$

donde $\eta > 0$, y x es un punto fijo en el que existen derivadas unilaterales $f'_+(x)$ y $f'_-(x)$.

Es obvio que la integral de Fourier

$$\frac{1}{\pi} \int_0^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\cos y(x-t)dt \qquad (56.7)$$

será el límite para la función (56.6) cuando $\eta \to +\infty$, es decir, $S(\eta)$ es en este sentido un análogo de las sumas parciales de las series de Fourier.

De conformidad con el teorema sobre la integración de integrales dependientes de un parámetro (véase el p. 53.1), para todo número $\xi > 0$ tenemos

$$\int_0^{\eta} dy \int_{-\xi}^{\xi} f(t)\cos y(x-t)dt = \int_{-\xi}^{\xi} f(t)dt \int_0^{\eta} \cos y(x-t)dy =$$

$$= \int_{-\xi}^{\xi} f(t) \frac{\operatorname{sen} \eta(x-t)}{x-t} dt. \qquad (56.8)$$

En efecto, ya que la función $f(t)$ es continua a trozos, sirviéndonos de las rectas paralelas al eje Oy, podemos dividir el rectángulo $-\xi \leqslant t \leqslant \xi$, $0 \leqslant y \leqslant \eta$, en un número finito de rectángulos, en cada uno de los cuales la función $f(t)\cos y(x-t)$ ya será continua, como función de dos variables, hasta la frontera (si en la frontera de los rectángulos mencionados por valores de la función f se toman, cuando sea necesario, sus límites unilaterales, es decir, $f(t+0)$ o bien $f(t-0)$). Al aplicar el teorema 3 del p. 53.1 a todo rectángulo y sumar los resultados obtenidos, obtendremos precisamente la fórmula (56.8).

De la desigualdad obvia

$$|f(t)\cos y(x-t)| \leqslant |f(t)|$$

y de la convergencia de la integral $\int_{-\infty}^{+\infty} |f(t)|\,dt$ proviene la convergencia uniforme, en el segmento $[0, \eta]$ respecto del parámetro y, de la integral

$$F(y) = \int_{-\infty}^{+\infty} f(t)\cos y(x-t)\,dt, \tag{56.9}$$

es decir,

$$F(y, \xi) = \int_{-\xi}^{\xi} f(t)\cos y(x-t)\,dt$$

tiende al límite (56.9), para $\xi \to +\infty$, uniformemente en el segmento $[0, \eta]$.

Luego, la función $F(y, \xi)$ es continua respecto de y. Efectivamente, la función f está acotada en el segmento $[-\xi, \xi]$: $|f(t)| \leqslant M$, $-\xi \leqslant t \leqslant \xi$. Designemos con $\omega(\delta)$ el módulo de continuidad de la función $\cos y(x-t)$, $0 \leqslant y \leqslant \eta$, $-\xi \leqslant t \leqslant \xi$. Entonces, $\lim_{\delta \to 0} \omega(\delta) = 0$, y, por eso

$$|F(y+\Delta y, \xi) - F(y, \xi)| \leqslant$$

$$\leqslant \int_{-\xi}^{\xi} |f(t)|\,|\cos(y+\Delta y)(x-t) - \cos y(x-t)|\,dt \leqslant 2M\xi\omega(\Delta y) \to 0$$

cuando $\Delta y \to 0$.

En virtud del teorema 2 del p. 53.1, en el primer miembro de la igualdad (56.8) podemos pasar al límite bajo el signo de la integral para $\xi \to +\infty$.

Como resultado tendremos

$$S(\eta) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t)\,\frac{\operatorname{sen}\eta(x-t)}{x-t}\,dt.$$

Esta integral es finita, pues (véanse (56.6) y (56.9)) es igual $\int_0^{\eta} F(y)\,dy$, donde la función $F(y)$ es continua como un límite, para $\xi \to +\infty$, de una familia de funciones $F(y, \xi)$ convergentes según y.

La integral $S(\eta)$ es un análogo de la integral de Dirichlet para las series de Fourier. Al poner $u = t - x$ (compárese con (55.17)), obtendremos

$$S(\eta) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(u + x) \frac{\operatorname{sen} \eta u}{u} \, du.$$

Representando la integral obtenida como una suma de dos integrales

$$\int_{-\infty}^{+\infty} = \int_{-\infty}^{0} + \int_{0}^{+\infty}$$

y realizando en la primera de ellas la sustitución $u = -t$, obtenemos

$$S(\eta) = \frac{1}{\pi} \int_{0}^{+\infty} [f(x + t) + f(x - t)] \frac{\operatorname{sen} \eta t}{t} \, dt.$$

Al recordar (véase el p. 54.4) que para $\eta > 0$

$$\int_{0}^{+\infty} \frac{\operatorname{sen} \eta t}{t} \, dt = \frac{\pi}{2},$$

obtenemos

$$S(\eta) - \frac{f(x + 0) + f(x - 0)}{2} = \frac{1}{\pi} \int_{0}^{+\infty} [f(x + t) + f(x - t)] \frac{\operatorname{sen} \eta t}{t} \, dt -$$

$$- [f(x + 0) + f(x - 0)] \frac{1}{\pi} \int_{0}^{+\infty} \frac{\operatorname{sen} \eta t}{t} \, dt =$$

$$= \frac{1}{\pi} \int_{0}^{+\infty} \frac{f(x + t) - f(x + 0)}{t} \operatorname{sen} \eta t \, dt +$$

$$+ \frac{1}{\pi} \int_{0}^{+\infty} \frac{f(x - t) - f(x - 0)}{t} \operatorname{sen} \eta t \, dt. \quad (56.10)$$

Consideremos, por ejemplo, la primera integral que figura en el segundo miembro de esta igualdad. Dividámosla en dos integrales:

$$\int_{0}^{+\infty} = \int_{0}^{1} + \int_{1}^{\infty}.$$

Por cuanto

$$\lim_{t \to +0} \frac{f(x + t) - f(x + 0)}{t} = f'_{+}(x),$$

entonces $\dfrac{f(x+t)-f(x+0)}{t}$ es una función continua a trozos de la variable t en el segmento $[0, 1]$, por lo cual, en virtud del teorema 2 del p. 55.2,

$$\lim_{\eta \to +\infty} \int_0^1 \frac{f(x+t)-f(x+0)}{t} \operatorname{sen} \eta t \, dt = 0. \tag{56.11}$$

La función $\dfrac{f(x+t)}{t}$ es también continua a trozos en cualquier segmento del semieje $t \geqslant 1$, y, como

$$\left|\frac{f(x+t)}{t}\right| \leqslant |f(x+t)|,$$

se tiene

$$\int_1^{+\infty} \left|\frac{f(x+t)}{t}\right| dt \leqslant \int_1^{+\infty} |f(x+t)|\, dt = \int_{x+1}^{+\infty} |f(s)|\, ds \leqslant \int_{-\infty}^{+\infty} |f(s)|\, ds \leqslant +\infty,$$

es decir, $\dfrac{f(x+t)}{t}$ es absolutamente integrable en este semieje, y, por lo tanto, en virtud del mismo teorema,

$$\lim_{\eta \to +\infty} \int_1^{+\infty} \frac{f(x+t)}{t} \operatorname{sen} \eta t \, dt = 0. \tag{56.12}$$

Por fin, de la convergencia de la integral $\displaystyle\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dt$ (véase el p. 33.6), al realizar el cambio de la variable $u = \eta t$, obtenemos

$$\lim_{\eta \to +\infty} \int_\eta^{+\infty} \frac{f(x+0)}{t} \operatorname{sen} \eta t \, dt = f(x+0) \lim_{\eta \to +\infty} \int_\eta^{+\infty} \frac{\operatorname{sen} u}{u}\, du = 0. \tag{56.13}$$

De (56.11), (56.12) y (56.13) se deduce que

$$\lim_{\eta \to +\infty} \frac{1}{\pi} \int_0^{+\infty} \frac{f(x+t)-f(x+0)}{t} \operatorname{sen} \eta t \, dt = 0.$$

De modo análogo se demuestra que

$$\lim_{\eta \to +\infty} \frac{1}{\pi} \int_0^{+\infty} \frac{f(x-t)-f(x-0)}{t} \operatorname{sen} \eta t \, dt = 0.$$

De aquí, en virtud de (56.10), obtendremos

$$\lim_{\eta \to +\infty} S(\eta) = \frac{f(x+0)+f(x-0)}{2}.$$

Por cuando el límite en el primer miembro es igual a la integral de Fourier (56.7), la igualdad (56.5) queda demostrada. □

Los requisitos impuestos sobre la función en este teorema pueden debilitarse al exigir, por ejemplo, que la función sea absolutamente integrable en todo el eje numérico y satisfaga en todo punto la condición generalizada de Hölder. No lo hemos hecho para hacer la demostración más simple (compárese con la demostración del teorema 4 y sus corolarios en el p. 55.4).

Ejercicio 1. Demuéstrese que si la función f es, en adición a las restricciones impuestas sobre ella en el teorema 1, par o impar, entonces quedan válidas las siguientes fórmulas: para una función par

$$\frac{f(x+0)+f(x-0)}{2}=\frac{2}{\pi}\int_0^{+\infty}\cos yx\,dy\int_0^{+\infty}f(t)\cos yt\,dt,$$

para una función impar

$$\frac{f(x+0)+f(x-0)}{2}=\frac{2}{\pi}\int_0^{+\infty}\operatorname{sen} yx\,dy\int_0^{+\infty}f(t)\operatorname{sen} yt\,dt.$$

56.2. DIFERENTES FORMAS DE NOTACIÓN DE LA FÓRMULA DE FOURIER

Para simplificar las notaciones se considerará en adelante que la función f es absolutamente integrable en todo el eje numérico R y en todos los puntos de éste es continua y tiene derivadas unilaterales. En este caso para cualquier $x \in R$, de acuerdo con el teorema 1, resulta lícita la fórmula de Fourier

$$f(x)=\frac{1}{\pi}\int_0^{+\infty}dy\int_{-\infty}^{+\infty}f(t)\cos y(x-t)\,dt,$$

y, por cuanto la función subintegral es par respecto de la variable y, entonces

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}dy\int_{-\infty}^{+\infty}f(t)\cos y(x-t)\,dt. \tag{56.14}$$

Por ser obvia la desigualdad

$$|f(t)\operatorname{sen} y(x-t)| \leqslant |f(t)|,$$

siendo vigentes las restricciones impuestas sobre la función f, existe también la integral

$$\int_{-\infty}^{+\infty}f(t)\operatorname{sen} y(x-t)\,dt,$$

con la particularidad de que, en virtud del criterio de Weierstrass (véase el p. 54.1), converge uniformemente en todo el eje numérico de la variable y y, por consiguien-

te, es una función continua de y. Por eso para cualquier número η existe una integral

$$\int_{-\eta}^{\eta} dy \int_{-\infty}^{+\infty} f(t)\operatorname{sen} y(x-t)\,dt, \tag{56.15}$$

y es, además, nula, puesto que la función subintegral es impar respecto de y. Sin embargo, asumidas las suposiciones referentes a la función f, no se puede garantizar la existencia de la integral impropia

$$\int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\operatorname{sen} y(x-t)\,dt. \tag{56.16}$$

Para obtener las fórmulas necesarias, hemos de introducir una generalización más del concepto de integral.

56.3. VALOR PRINCIPAL DE UNA INTEGRAL

Introduzcamos la siguiente definición.

Definición 2. *Sea φ una función integrable en cualquier segmento finito. Si existe un límite finito*

$$\lim_{\eta \to +\infty} \int_{-\eta}^{\eta} \varphi(x)\,dx, \quad \eta > 0,$$

se denominará valor principal de la integral $\int_{-\infty}^{+\infty} \varphi(x)\,dx$ *y se designará brevemente* v.p.

$$\text{v.p.} \int_{-\infty}^{+\infty} \varphi(x)\,dx \stackrel{\text{def}}{=} \lim_{\eta \to +\infty} \int_{-\eta}^{\eta} \varphi(x)\,dx. \tag{56.17}$$

Subrayemos que la diferencia de esta definición de la que caracteriza una integral impropia $\int_{-\infty}^{+\infty} \varphi(x)\,dx$, en el sentido de la definición dada en el p. 33.1, consiste en que para la función φ, integrable en cualquier segmento finito, la integral $\int_{-\infty}^{+\infty} \varphi(x)\,dx$ se definía como un límite de las integrales $\int_{-\infty}^{+\infty} \varphi(x)\,dx$, cuando tendían independientemente ξ hacia $-\infty$ y η hacia $+\infty$. En el caso que se considera ahora sólo se exige la existencia del límite de las integrales citadas $\int_{\xi}^{\eta} \varphi(x)\,dx$ para un caso particular en que $\xi = -\eta$ y $\eta \to +\infty$.

De este mismo modo se determina también el valor principal de la integral impropia en un punto: supongamos que $a < c < b$ y la función φ es integrable según Riemann, para cualquier $\varepsilon > 0$, en los segmentos $[a, c-\varepsilon]$ y $[c+\varepsilon, b]$ (se supone, naturalmente, que $a < c - \varepsilon$ y $c + \varepsilon < b$); entonces el valor principal de la

integral $\int_a^b \varphi(x)dx$ en el punto c se determinará mediante la fórmula

$$v.p. \int_a^b \varphi(x)dx = \lim_{\varepsilon \to 0} \left[\int_a^{c-\varepsilon} \varphi(x)dx + \int_{c+\varepsilon}^b \varphi(x)dx \right].$$

A veces, cuando ello no pueda llevar a las equivocaciones, la integral en el sentido del valor principal se designa simplemente con el símbolo integral, omitiéndose las letras *v.p.*

Si para una cierta función existe la integral impropia, dicha función dispone también del valor principal de la integral y éste coincide con la integral impropia de la función. Lo recíproco no es cierto: para una función puede existir (y, por consiguiente, ser finito) el valor principal de la integral, mientras que la integral impropia es divergente.

Por ejemplo, las integrales $\int_{-\infty}^{+\infty} x\,dx$ y $\int_{-1}^{1} \frac{dx}{x}$ no existen como impropias, no obstante existen en el sentido del valor principal, el cual en ambos casos es igual a cero.

56.4. NOTACIÓN COMPLEJA DE LA INTEGRAL DE FOURIER

Volveremos a la fórmula de Fourier (56.14) y la escribiremos en otra forma, sirviéndonos del concepto de valor principal de la integral. Puesto que la función subintegral en la integral (56.16) es impar respecto de y, con arreglo a la definición enunciada del valor principal de la integral, tenemos

$$v.p. \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t) \operatorname{sen} y(x-t)dt = 0. \tag{56.18}$$

Al multiplicar ambos miembros de esta igualdad por $\frac{i}{2\pi}$ y sumar con la integral (56.14), obtendremos

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t)e^{iy(x-t)}dt, \tag{56.19}$$

donde la integral exterior se entiende en el sentido del valor principal. La fórmula (56.19) lleva el nombre de *notación compleja de la integral de Fourier.*

56.5. TRANSFORMACIÓN DE FOURIER

Si ponemos

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{-iyt}dt,$$

la fórmula (56.19) tendrá por expresión

$$f(x) = v.p.\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Phi(y) e^{ixy}\, dy. \tag{56.20}$$

Definición 3. *La función Φ que se pone en correspondencia con la función f por la fórmula*

$$\Phi(y) = v.p.\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-iyt}\, dt, \tag{56.21}$$

se denomina transformación de Fourier de la función f y se designa $F[f]$ o bien f.

En esta definición $f(t)$ es, en el caso general, una función de valores complejos del argumento real. Notemos que la función $\Phi = F[f]$ puede adquirir valores esencialmente complejos incluso cuando la función f toma sólo valores reales.

La transformación de Fourier está definida, en particular, para todas las funciones absolutamente integrables. Por ejemplo, empleando para la transformación de Fourier de la función f la designación $\hat{f}$, podemos escribir la fórmula (56.20) en la forma

$$f(x) = v.p.\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(y) e^{ixy}\, dy. \tag{56.20'}$$

Esta fórmula permite restablecer la propia función f, si se sabe su transformación de Fourier f. Se denomina *fórmula de inversión.*

Definición 4. *La función Ψ que se pone en correspondencia con la función f mediante la fórmula*

$$\Psi(y) = v.p.\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{iyt}\, dt, \tag{56.22}$$

se llama transformación inversa de Fourier de la función f y designa por $F^{-1}[f]$.

La transformación de Fourier y la transformación inversa de Fourier están definidas en un conjunto de funciones, para las cuales las integrales (56.21) y (56.22) existen en el sentido del valor principal. Este conjunto contiene dentro de sí, en particular, el conjunto de todas las funciones absolutamente integrables en todo el eje real, para las cuales las integrales en las fórmulas (56.21) y (56.22) pueden entenderse como integrales impropias ordinarias y no sólo como integrales en el sentido del valor principal. El término "transformación inversa de Fourier" se justifica por lo que la transformación F^{-1} convierte la transformación de Fourier F. Con más precisión, es válido el siguiente lema.

Lema 1. *Si una función f, continua y absolutamente integrable en todo el eje, tiene en cada punto derivadas unilaterales finitas, entonces*

$$F^{-1}[F[f]] = F[F^{-1}[f]] = f.$$

DEMOSTRACIÓN. La primera fórmula de inversión, es decir, la fórmula $F^{-1}[F[f]] = f$ es simplemente otra notación de la fórmula ya demostrada (56.19).

Probemos la validez de la segunda fórmula de inversión. Por cuanto el coseno es una función par, podemos en (56.14) cambiar de lugares t y x;

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\cos y(t - x)\,dt,$$

y, por ser impar el seno (compárese con (56.18)),

$$v.p. \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t)\operatorname{sen} y(t - x)\,dt = 0.$$

Por eso, a la par con la fórmula (56.19), tenemos también

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(t)e^{iy(t-x)}\,dt,$$

o bien

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{ity}\,dt\right] e^{ixy}\,dy,$$

donde la integral exterior se toma en el sentido del valor principal. Esta fórmula puede ser escrita en la forma

$$F[F^{-1}[f]] = f. \quad \square$$

Indiquemos que la validez de las fórmulas de inversión puede demostrarse también para restricciones más débiles impuestas sobre la función en comparación con aquellas que prevén la existencia en cada punto de las derivadas unilaterales.

Lema 2. *Supongamos que para las funciones f_1 y f_2 existe la transformación de Fourier (la transformación inversa de Fourier, respectivamente). Entonces, cualesquiera que sean los números λ_1 y λ_2, para las funciones $\lambda_1 f_1 + \lambda_2 f_2$ también existe la transformación de Fourier (la transformación inversa de Fourier, respectivamente), con la particularidad de que*

$$F[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1 F[f_1] + \lambda_2 F[f_2]$$

$(F^{-1}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1 F^{-1}[f_1] + \lambda_2 F^{-1}[f_2]$, *respectivamente*).

Esta propiedad lleva el nombre de *linealidad de la transformación de Fourier* (de la transformación inversa de Fourier, respectivamente). Ella proviene inmediatamente de la linealidad de la integral y de las fórmulas (56.21) y (56.22).

Corolario. $F[0] = F^{-1}[0] = 0$.

En efecto, por ejemplo,

$$F[0] = F[0 \cdot 0] = 0 \cdot F[0] = 0.$$

Esta última propiedad se deduce, naturalmente, asimismo de las fórmulas (56.21) y (56.22).

Lema 3. *La transformación de Fourier F, al igual que la transformación inversa de Fourier F^{-1}, son las aplicaciones biunívocas de un conjunto de funciones continuas absolutamente integrables en todo el eje real, que tienen en cada punto derivadas unilaterales, en otro conjunto de funciones, para las cuales las integrables* (56.21) *y* (56.22) *existen en el sentido del valor principal.*

DEMOSTRACIÓN. Es suficiente demostrar sólo la biunivocidad de las aplicaciones F y F^{-1}, puesto que lo demás ya ha sido demostrado más arriba. Demostremos, por ejemplo, la biunivocidad de la aplicación F. Sea $F[f_1] = F[f_2]$; entonces

$$F^{-1}[F[f_1]] = F^{-1}[F[f_2]].$$

De aquí, de acuerdo con el lema 1, se desprende que

$$f_1 = f_2. \quad \square$$

En todo caso la transformación de Fourier está definida para las funciones absolutamente integrables. En los puntos que siguen se estudiarán las propiedades de esta transformación. Más adelante aún se mostrará como la transformación de Fourier se generaliza a las clases más amplias de funciones, a saber, a las funciones de cuadrado integrable (el p. 58.7*) y las así llamadas funciones generalizadas.

56.6. INTEGRALES DE LAPLACE

Hallemos la transformación de Fourier $\hat{f}$ de la prolongación par de la función e^{-ax}, $a > 0$, desde la semirecta $x \geqslant 0$ a toda la recta numérica, es decir, hablando sencillamente, la transformación de Fourier de la función $f(x) = e^{-a|x|}$, $-\infty < x < +\infty$:

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-a|x|} e^{-ixy} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{ax} e^{-ixy} dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-ax} e^{-ixy} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-(a-iy)x} dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-(a+iy)x} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{a - iy} + \frac{1}{a + iy} \right) = \sqrt{\frac{2}{\pi}} \frac{a}{a^2 + y^2}.$$

La aplicación de la transformación inversa de Fourier a la función obtenida da la función de partida

$$e^{-ax} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sqrt{\frac{2}{\pi}} \frac{a}{a^2 + y^2} e^{ixy} dy, \quad x \geqslant 0.$$

Recordando que $e^{ixy} = \cos xy + i \operatorname{sen} xy$ y notando que, en virtud de que la función subintegral de la integral $\int_{-\infty}^{+\infty} \frac{\operatorname{sen} xy}{a^2 + y^2}\, dy = 0$ obtendremos

$$e^{-ax} = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{\cos xy}{a^2 + y^2}\, dy = \frac{2a}{\pi} \int_{0}^{+\infty} \frac{\cos xy}{a^2 + y^2}\, dy, \; x \geqslant 0.$$

Hallemos ahora la transformación de Fourier $\hat{f}$ de la prolongación impar de la función e^{-ax}, $a > 0$, desde el semieje positivo $x > 0$, es decir, la transformación de Fourier de la función

$$f(x) = \begin{cases} e^{-ax} & x > 0, \\ -e^{ax} & x < 0. \end{cases}$$

Tenemos

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} (-e^{ax}) e^{-ixy}\, dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-ax} e^{ixy}\, dx =$$

$$= \frac{1}{\sqrt{2\pi}} \left(-\frac{1}{a - iy} + \frac{1}{a + iy} \right) = -\sqrt{\frac{2}{\pi}} \frac{y}{a^2 + y^2}\, i.$$

Al aplicar nuevamente la fórmula de inversión de la transformación de Fourier, obtendremos

$$e^{-ax} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(-\sqrt{\frac{2}{\pi}} \frac{y}{a^2 + y^2}\, i \right) e^{ixy}\, dy = \frac{2}{\pi} \int_{0}^{+\infty} \frac{y \operatorname{sen} xy}{a^2 + y^2}\, dy, \quad x > 0.$$

Así pues, hemos logrado no sólo hallar la transformación de Fourier de las funciones en consideración, sino también obtener directamente de la fórmula de inversión (56.20') los valores de dos integrales

$$\int_{0}^{+\infty} \frac{\cos xy}{a^2 + y^2}\, dy = \frac{\pi}{2a} e^{-ax}, \quad x \geqslant 0,$$

$$\int_{0}^{+\infty} \frac{y \operatorname{sen} xy}{a^2 + y^2}\, dy = \frac{\pi}{2} e^{-ax}, \quad x > 0.$$

Estas últimas se llaman *integrales de Laplace*.

56.7. PROPIEDADES DE LAS TRANSFORMACIONES DE FOURIER DE LAS FUNCIONES ABSOLUTAMENTE INTEGRABLES

En este punto y en los que siguen serán consideradas algunas propiedades de la transformación de Fourier de la función f, la cual se designará, como hasta ahora,

mediante $\hat{f}$ o $F[f]$. Se supondrá, además, que la función f toma, en el caso general, los valores complejos, mientras que su argumento es, como siempre, real.

Lema 4. *Si la función f es absolutamente integrable en todo el eje numérico, su transformación de Fourier $\hat{f}(y)$ está acotada en todo el eje, con la particularidad de que*

$$|\hat{f}(y)| \leqslant \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |f(x)|\,dx. \tag{56.23}$$

Corolario. *Si una sucesión de las funciones absolutamente integrables $f_n(x)$, $n = 1, 2, \ldots$, y una función absolutamente integrable $f(x)$ son de tal género que*

$$\lim_{n \to \infty} \int_{-\infty}^{+\infty} |f_n(x) - f(x)|\,dx = 0,$$

entonces la sucesión $\{\hat{f}_n(y)\}$ converge uniformemente en todo el eje numérico hacia la función $\hat{f}(y)$.

DEMOSTRACIÓN. La desigualdad (56.23) se infiere de la fórmula (véase (56.21))

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ixy} dy, \tag{56.24}$$

si se tiene presente que $|e^{-ixy}| = 1$. □

El corolario se deduce inmediatamente de la desigualdad (56.23) y la linealidad de la transformación de Fourier, pues

$$|\hat{f}_n(y) - \hat{f}(y)| = |\widehat{f_n(x) - f(x)}| \underset{(56.23)}{\leqslant} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |f_n(x) - f(x)|\,dx. \quad \square$$

Lema 5. *Si la función f es absolutamente integrable en todo el eje numérico, su transformación de Fourier $\hat{f}(y)$ es continua y*

$$\lim_{y \to \pm\infty} \hat{f}(y) = 0. \tag{56.25}$$

DEMOSTRACIÓN. Sea $f(x) = u(x) + iv(x)$, donde $u(x)$ y $v(x)$ son unas funciones reales absolutamente integrables. Por cuanto $\widehat{f(x)} = \widehat{u(x)} + i\widehat{v(x)}$, para demostrar la continuidad de la función $\hat{f}(y)$ es suficiente demostrar la continuidad de las funciones $\widehat{u(x)}$ y $\widehat{v(x)}$. Aquí $u(x)$ y $v(x)$ son, como siempre, las funciones de valores reales de un argumento real, mientras que $\widehat{u(x)}$ y $\widehat{v(x)}$ son, en el caso general, funciones de valores complejos de un argumento real.

De acuerdo con el lema 2 del p. 55.2, para cualquier función $f(x)$, que es real y absolutamente integrable en todo el eje, existe una sucesión de funciones escalonadas finitas $\varphi_n(x)$, $n = 1, 2, \ldots$, tales que

$$\lim_{n \to \infty} \int_{-\infty}^{+\infty} |\varphi_n(x) - f(x)|\,dx = 0.$$

En virtud del corolario 4, la sucesión $[\hat{\varphi}_n(y)]$ converge uniformemente hacia la función $\hat{f}(y)$. Para convencerse de que la función $\hat{f}(y)$ es continua, es suficiente demostrar que las funciones $\hat{\varphi}_n(y)$ son continuas (véase el teorema 8′ en el p. 36.4). Mostrémoslo. Toda función escalonada finita es una combinación lineal de funciones de un solo escalón (véase el p. 55.2), con más precisión, de funciones características de los semiintervalos de tipo $[a, b)$. Por eso, debido a la linealidad de la transformación de Fourier, la continuidad de la función $\hat{f}$ será demostrada, si probamos que para la función característica de cualquier semiintervalo $[a, b)$ su transformación de Fourier es continua.

Sea ω una función característica del semiintervalo $[a, b)$, es decir, $\omega(x) = 1$, si $a \leqslant x < b$, y $\omega(x) = 0$, si $x < a$ ó $x \geqslant b$. Entonces, en virtud de (56.21), para $y \neq 0$, tenemos

$$\hat{\omega}(y) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-ixy} dx = -\frac{1}{iy\sqrt{2\pi}} \int_a^b e^{-ixy} d(-ixy) =$$

$$= \frac{i}{y\sqrt{2\pi}} e^{-ixy}\Big|_{x=a}^{x=b} = \frac{i(e^{-iby} - e^{-iay})}{y\sqrt{2\pi}}.$$

En el caso de $y = 0$, en vista de la misma fórmula (56.21),

$$\hat{\omega}(0) = \frac{1}{\sqrt{2\pi}} \int_a^b dx = \frac{b-a}{\sqrt{2\pi}}.$$

Así pues,

$$\hat{\omega}(y) = \begin{cases} \dfrac{i(e^{-iby} - e^{-iay})}{y\sqrt{2\pi}} & \text{para } y \neq 0, \\ \dfrac{b-a}{\sqrt{2\pi}} & \text{para } y = 0. \end{cases}$$

Es evidente que el segundo miembro de esta igualdad es una función continua para cualquier $y \neq 0$. Probemos que es también continua cuando $y = 0$.

$$\lim_{y \to 0} \frac{i(e^{-iyb} - e^{-iya})}{y\sqrt{2\pi}} = \frac{i}{\sqrt{2\pi}} \lim_{y \to 0} \frac{1}{y} [(1 - iby + o(y)) -$$

$$- (1 - iay + o(y))] = \frac{1}{\sqrt{2\pi}} \lim_{y \to 0} \left[b - a + \frac{o(y)}{y} \right] = \frac{b-a}{\sqrt{2\pi}},$$

es decir: la función $\hat{\omega}(y)$ es realmente continua en el punto $y = 0$.

De este modo queda demostrada la continuidad en todo el eje numérico de la transformación de Fourier $\hat{f}$ de una función f absolutamente integrable en todo el eje numérico que toma valores reales. De aquí, según lo dicho anteriormente, proviene inmediatamente la continuidad de la transformación de Fourier $\hat{f}$ de la función $f = u + iv$ absolutamente integrable en todo el eje numérico, es decir, de una función que, en el caso general, toma también valores complejos.

La igualdad (56.25) se infiere del teorema 2, p. 55.2. En efecto, supongamos nuevamente al principio que la función f es absolutamente integrable en todo el eje numérico y toma sólo valores reales, entonces

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}}\left[\int_{-\infty}^{+\infty} f(x)\cos xy\,dx - i\int_{-\infty}^{+\infty} f(x)\operatorname{sen} xy\,dx\right],$$

donde, en virtud del teorema citado, las partes real e imaginaria y, por consiguiente, la propia función $\hat{f}(y)$ tienden a cero cuando $y \to \pm\infty$.

Ahora, si $f = u + iv$, entonces, según lo demostrado $\lim_{y\to\pm\infty} \hat{u}(y) = \lim_{y\to\pm\infty} \hat{v}(y) = 0$, por consiguiente, $\lim_{y\to\pm\infty} \hat{f}(y) = 0$.

56.8. TRANSFORMACIÓN DE FOURIER DE LAS DERIVADAS

Teorema 2. *Supongamos que una función f, absolutamente integrable en todo el eje numérico, tiene n derivadas absolutamente integrables y continuas en todo el eje. Entonces,*

$$F[f^{(k)}] = (iy)^k F[f], \quad k = 0, 1, \ldots, n, \tag{56.26}$$

y existe una constante $M > 0$ tal que

$$|F[f]| \leqslant \frac{M}{|y^n|}. \tag{56.27}$$

DEMOSTRACIÓN. Supongamos primero que la función f admite sólo valores reales. Si f es absolutamente integrable en todo el eje junto con su derivada f' y esta derivada es continua, entonces

$$f(x) = f(0) + \int_0^x f'(t)\,dt.$$

Dado que la integral $\int_{-\infty}^{+\infty} |f'(t)|\,dt$ converge de conformidad con la hipótesis del teorema, entonces converge también la integral $\int_{-\infty}^{+\infty} f'(t)\,dt$, por lo cual, en virtud de la definición de convergencia de una integral, existen los límites $\lim_{x\to\pm\infty} \int_0^x f'(t)\,dt$ y, por consiguiente, los límites $\lim_{x\to\pm\infty} f(x)$. Además, de la convergencia de la integral $\int_{-\infty}^{+\infty} f(t)\,dt$ se deduce que los límites citados son iguales a cero: $\lim_{x\to\pm\infty} f(x) = 0$. Al aplicar la integración por partes a la fórmula de la transformación de Fourier, obtendremos:

$$F[f'] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(x)e^{-ixy}\,dx =$$

$$= \frac{1}{\sqrt{2\pi}} f(x)e^{-ixy}\Big|_{-\infty}^{+\infty} + \frac{iy}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)e^{-ixy}dx = iyF[f].$$

De este modo, la derivación de una función conduce a la multiplicación de su transformación de Fourier por el factor iy.

Si ahora $f = u + iv$, donde u y v son funciones reales, y esta vez de nuevo f es absolutamente integrable junto con su derivada $f' = u' + iv'$ y esta última es continua, entonces

$$F[f'] = F[u' + iv'] = F[u'] + iF[v'] = iyF[u] - yF[v] =$$

$$= iyF[u + iv] = iyF[f].$$

La fórmula (56.26) queda demostrada para $n = 1$. Para n arbitrario ésta se obtiene por inducción.

La función $F[f^{(n)}]$ está acotada (véase el lema 4), por lo cual la cota superior $M = \sup_{-\infty < y < +\infty} F[f^{(n)}]$ es finita y, por consiguiente, la estimación (56.27) se desprende de la fórmula (56.26) para $k = n$. □

Así pues, cuanto mayor es el número de las derivadas absolutamente integrables que tiene la función f, tanto más rápido tiende a cero en el infinito su transformación de Fourier.

Ha de notarse que el teorema 2, al igual que su demostración, quedan en vigor también en el caso cuando la derivada de n-ésimo orden de la función en consideración no es continua, sino que tiene un número finito de discontinuidades de primera especie (véase el p. 5.1), conservándose sin cambios otras suposiciones. En efecto, en el caso dado la derivada citada es, en cualquier segmento finito, una función continua a trozos (véase el p. 28.3) y por esta razón la integración por partes, a la que se recurre en la demostración, es lícita (véanse los pp. 30.2 y 33.2)

Ejercicio 2. Demuéstrese que la transformación de Fourier $F(y)$ de la función $f(x) = \frac{1}{1 + |x|^3}$ es igual a $O\left(\frac{1}{y^3}\right)$ cuando $y \to \infty$.

56.9. CONVOLUCIÓN Y TRANSFORMACIÓN DE FOURIER

Supongamos que las funciones φ y ψ están definidas en todo el eje real. En diferentes problemas matemáticos se emplea a menudo la así llamada *convolución de las funciones* φ y ψ, la que se designa por $\varphi * \psi$ o, si x es el argumento de la convolución, mediante $(\varphi * \psi)(x)$ y se determina con la igualdad

$$(\varphi * \psi)(x) = \int_{-\infty}^{+\infty} \varphi(t)\psi(x - t)dt. \tag{56.28}$$

En este punto supondremos, para simplificar, que las funciones en consideración $\varphi(t)$, $\psi(t)$, $\chi(t)$ admiten sólo los valores reales. La integral (56.28) existe a cien-

cia cierta, si ambas funciones están acotadas y son absolutamente integrables *). Además, la integral (56.28) y, más aún, la integral

$$\int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dt$$

son uniformemente convergentes en todo el eje real. Efectivamente, en virtud de que la función ψ está acotada se tiene $|\psi| \leqslant M$, donde M es una constante, por lo cual para cualquiera x y t

$$|\varphi(t)\psi(x-t)| \leqslant M|\varphi(t)|$$

y la afirmación enunciada se deduce, en virtud de la integrabilidad absoluta de la función φ, del criterio de Weierstrass para la convergencia uniforme de integrales (véase el p. 54.1). De los razonamientos aducidos se desprende, además, que si las funciones φ y ψ están acotadas y son absolutamente integrables y continuas, su convolución f es también continua, acotada y absolutamente integrable. En efecto, la continuidad de la función f proviene de la convergencia uniforme de la integral (56.28) y el carácter acotado, de la estimación

$$|(\varphi * \psi)(x)| \leqslant \int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dt \leqslant M \int_{-\infty}^{+\infty} |\varphi(t)|\,dt.$$

Demostremos la integrabilidad absoluta de la convolución. Sea $f = \varphi * \psi$; tenemos

$$\int_{-\infty}^{+\infty} |f(x)|\,dx = \int_{-\infty}^{+\infty} dx \left| \int_{-\infty}^{+\infty} \varphi(t)\psi(x-t)\,dt \right| \leqslant$$

$$\leqslant \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dt = \int_{-\infty}^{+\infty} |\varphi(t)|\,dt \int_{-\infty}^{+\infty} |\psi(x-t)|\,dx =$$

$$= \int_{-\infty}^{+\infty} |\varphi(t)|\,dt \int_{-\infty}^{+\infty} |\psi(s)|\,ds. \quad (56.29)$$

El cambio del orden de integración en este caso es posible en virtud de que (véase el teorema 7 del p. 54.3) la integral $\int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dt$ converge uniformemente en todo el eje, la integral $\int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dx = |\varphi(t)| \int_{-\infty}^{+\infty} |\psi(x-t)|\,dx$ converge uniformemente en cualquier segmento finito (¿por qué?), y la integral reiterada $\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} |\varphi(t)\psi(x-t)|\,dt$ existe, según se pone de manifiesto de la última igualdad de la fórmula (56.29).

*) La existencia de la integral (56.28) puede garantizarse también para las condiciones más generales, sin embargo no nos detendremos en esto.

De este modo, asumidas las suposiciones especificadas, podemos aplicar a la función $f = \varphi * \psi$ la operación de convolución con una función continua acotada y absolutamente integrable (como resultado, se obtendrá nuevamente una función de la misma clase) o la transformación de Fourier.

La operación de convolución de las funciones es *conmutativa y asociativa* en la clase de funciones en consideración. En efecto, al realizar en la integral (56.28) el cambio de la variable $x - t = s$, obtendremos

$$\varphi * \psi = \int_{-\infty}^{+\infty} \varphi(t)\psi(x - t)dt = \int_{-\infty}^{+\infty} \varphi(x - s)\psi(s)ds = \psi * \varphi$$

Luego, realizando en la integral que viene abajo el cambio de la variable $t = y - \xi$, cambiando el orden de integración y haciendo la sustitución $x - y + \xi = \eta$, obtendremos

$$(\varphi * \psi) * \chi = \int_{-\infty}^{+\infty} \chi(y - x)dx \int_{-\infty}^{+\infty} \varphi(t)\psi(x - t)dt =$$

$$= \int_{-\infty}^{+\infty} \chi(y - x)dx \int_{-\infty}^{+\infty} \varphi(y - \xi)\psi(x - y + \xi)d\xi =$$

$$= \int_{-\infty}^{+\infty} \varphi(y - \xi)d\xi \int_{-\infty}^{+\infty} \psi(x - y + \xi)\chi(y - x)dx =$$

$$= \int_{-\infty}^{+\infty} \varphi(y - \xi)d\xi \int_{-\infty}^{+\infty} \psi(\eta)\chi(\xi - \eta)d\eta = (\psi * \chi) * \varphi = \varphi * (\psi * \chi).$$

La posibilidad de cambiar el orden de integración en este caso también se desprende del teorema 7d el p. 54.3. En efecto, investiguemos la convergencia uniforme de las integrales

$$\chi(y - x) \int_{-\infty}^{+\infty} \varphi(y - \xi)\psi(x - y + \xi)d\xi, \tag{56.30}$$

$$\varphi(y - \xi) \int_{-\infty}^{+\infty} \psi(x - y + \xi)\chi(y - x)dx. \tag{56.31}$$

Puesto que las funciones ψ y χ son acotadas, se tiene $|\psi| \leqslant M$, $|\chi| \leqslant M$, donde M es una constante, por lo cual

$$|\chi(y - x)\varphi(y - \xi)\psi(x - x + \xi)| \leqslant M^2|\varphi(y - \xi)|,$$

$$|\varphi(y - \xi)\psi(x - y + \xi)\chi(y - x)| \leqslant M^2|\chi(y - x)|.$$

De estas igualdades y del hecho de que las funciones φ y χ son absolutamente integrables se deduce, de acuerdo con el criterio de Weierstrass, que las integrales (56.30) y (56.31) convergen uniformemente respecto de las variables x y ξ, respectivamente (la variable y está fija) en cualquier segmento finito (¿por qué?). Por fin, existe una integral reiterada

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} |\chi(y - x)\varphi(y - \xi)\psi(x - y + \xi)|d\xi = (|\varphi| * |\psi|) * |\chi|.$$

De este modo, todas las condiciones del teorema mencionado 7 del p. 54.3 están cumplidas.

Cabe notar que al considerar las convoluciones de las funciones, se pueden debilitar considerablemente las restricciones que se imponen sobre las funciones de convolución. No obstante, la demostración de las propiedades de convoluciones en este caso exigiría, ante todo, unos teoremas más delicados sobre el cambio del orden de integración. No lo hicimos aquí con el fin de simplificar la exposición.

Procedamos ahora al estudio de la transformación de Fourier de la convolución de dos funciones. Transformemos, para la comodidad, la definición de la convolución $\varphi * \psi$, añadiendo el factor adicional $1/\sqrt{2\pi}$:

$$\varphi * \psi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(t)\psi(x - t)dt.$$

Teorema 3. *Sean φ y ψ las funciones acotadas, continuas y absolutamente integrables en todo el eje numérico. En este caso tenemos*

$$F[\varphi * \psi] = F[\varphi]F[\psi].$$

DEMOSTRACIÓN. Las funciones φ y ψ son acotadas, continuas y absolutamente integrables y por ello la función $\varphi * \psi$ posee las mismas propiedades, en particular, es absolutamente integrable y para ella puede considerarse la transformación de Fourier

$$F[\varphi * \psi] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ixy}dx \int_{-\infty}^{+\infty} \varphi(t)\psi(x - t)dt.$$

Cambiando aquí el orden de integración (lo que es posible en virtud del teorema 7 del p. 54.3) y realizando el cambio de la variable $x = t + s$, obtendremos

$$F[\varphi * \psi] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \varphi(t)dt \int_{-\infty}^{+\infty} \psi(x - t)e^{-ixy}dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(t)e^{-ity}dt \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(s)e^{-isy}ds = F[\varphi]F[\psi],$$

es decir, la transformación de Fourier de una convolución de dos funciones es igual al producto de las transformaciones de Fourier de dichas funciones. □

El teorema 3 puede ser demostrado también para restricciones más débiles impuestas sobre las funciones en consideración, pero no nos detendremos en esto.

56.10. DERIVADA DE LA TRANSFORMACIÓN DE FOURIER DE UNA FUNCIÓN

Teorema 4. *Si la función $f(x)$ es continua y las funciones $f(x)$, $xf(x)$, . . . , $x^n f(x)$ son absolutamente integrables en todo el eje numérico , entonces la transfor-*

mación de Fourier de la función f será una función n veces derivable en todo el eje numérico y

$$i^k F^{(k)}[f] = F[x^k f], \quad k = 0, 1, \ldots, n.$$

DEMOSTRACIÓN. Supongamos primero que la función f toma sólo valores reales. Derivando formalmente la integral

$$F[f] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ixy} dx$$

respecto del parámetro y y observando que $|xf(x)e^{-ixy}| = |xf(x)|$, obtendremos la integral absoluta e uniformemente convergente

$$-i \int_{-\infty}^{+\infty} xf(x) e^{-ixy} dx, \quad -\infty < y < +\infty.$$

Por consiguiente (véase el p. 54.3, teorema 8) en este caso la transformación de Fourier $F[f]$ de la función f es una función derivable e

$$iF'[f] = F[xf].$$

Si, ahora, $f = u + iv$, donde u y v son unas funciones reales obtenemos

$$F'[f] = F'[u + iv] = \{F[u] + iF[v]\}' = F'[u] + iF'[v] =$$
$$= -iF[xu] + F[xv] = -iF[xu + ixv] = -iF[xf].$$

A continuación, por inducción obtenemos que la transformación de Fourier $F[f]$ de la función f tiene derivadas hasta el orden n inclusive e $i^k F^{(k)}[f] = F[x^k f]$, $k = 0, 1, \ldots, n$. □

Corolario. *Si las suposiciones del teorema están cumplidas, todas las derivadas $F^{(k)}[f]$, $k = 0, 1, \ldots, n$, son continuas y tienden a cero cuando su argumento tiende al infinito.*

En vista del lema 5, el corolario se deduce inmediatamente, ya que las derivadas $F^{(k)}[f]$ son las transformaciones de Fourier de las funciones absolutamente integrables.

Se puede mostrar que si los productos del tipo $e^{a|x|^\alpha} f(x)$ son absolutamente integrables, asumidas ciertas restricciones impuestas sobre $a > 0$ y $\alpha > 0$, esto conduce a que la transformación de Fourier se hace aún más suave, a saber, la transformación mencionada pertenece ya a unas u otras clases de funciones analíticas.

La fórmula que determina la transformación inversa de Fourier se diferencia de la que da la transformación directa de Fourier (véanse (56.21) y (56.22)) sólo en que el número e bajo el signo integral tiene i sustituido por $-i$ en el exponente de la potencia, razón por la cual para la transformación inversa de Fourier son válidas las propiedades análogas a las demostradas para la transformación directa de Fourier.

Ejercicios. 3. Demuéstrese que la transformación de Fourier de la función $f(x) = \dfrac{1}{1 + x^4}$ es dos veces derivable en todo el eje numérico.

4. Demuéstrese que la transformación de Fourier de la función $f(x) = xe^{-|x|}$ es infinitamente derivable en todo el eje numérico.

§ 57. ESPACIOS FUNCIONALES

57.1. ESPACIOS MÉTRICOS

Definición 1. *El conjunto $X = \{x, y, z, \ldots\}$ se llama espacio métrico X, si en un conjunto de pares ordenados (x, y) de elementos de este conjunto viene definida una función no negativa $\rho(x, y)$, llamada distancia (o métrica), tal que:*

1) $\rho(x, y) = 0$ *cuando, y sólo cuando,* $x = y$;

2) $\rho(x, y) = \rho(y, x)$, $x \in X$, $y \in X$;

3) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$, $x \in X$, $y \in X$, $z \in X$.

Las condiciones 1, 2 y 3 se denominan *axiomas de distancia.*

Los elementos del espacio métrico llevan el nombre de *puntos.*

Ejemplos. 1. Una totalidad de todos los números reales R forma un espacio métrico, siempre que la distancia entre los números reales se define como valor absoluto de la diferencia entre ellos:

$$\rho(x, y) = |x - y|, \quad x \in R, \quad y \in R.$$

2. Un conjunto de todos los números complejos C forma también un espacio métrico, si la distancia entre los elementos del conjunto se define según la fórmula $\rho(z, z') = |z - z'|$, $z \in C$, $z' \in C$.

3. Un espacio euclídeo R^n de dimensión n (véase el p. 18.1) es un espacio métrico, si la distancia entre sus puntos $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ se determina por la fórmula (véase (18.1))

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

4. Sea X cierto conjunto. Consideraremos un conjunto de funciones numéricas que son acotadas en X y que toman valores reales (o complejos). Para dos funciones de esta índole φ y ψ pongamos

$$\rho(\varphi, \psi) = \sup_{t \in X} |\varphi(t) - \psi(t)|. \tag{57.1}$$

Se comprueba con facilidad que la función $\rho(\varphi, \psi)$ es una métrica. La validez de las propiedades 1 y 2 de la distancia se ve claramente. Comprobemos la validez de la propiedad 3. Supongamos que φ, ψ y χ son unas funciones acotadas definidas en el conjunto X. Para todo elemento $t \in X$ tenemos

$$|\varphi(t) - \chi(t)| = |[\varphi(t) - \psi(t)] + [\psi(t) - \chi(t)]| \leqslant$$
$$\leqslant |\varphi(t) - \psi(t)| + |\psi(t) - \chi(t)|,$$

por lo cual

$$|\varphi(t) - \chi(t)| \leqslant \sup_X |\varphi(t) - \psi(t)| + \sup_X |\psi(t) - \chi(t)|,$$

de donde

$$\sup_X |\varphi(t) - \chi(t)| \leqslant \sup_X |\varphi(t) - \psi(t)| + \sup_X |\psi(t) - \chi(t)|,$$

es decir,

$$\rho(\varphi, \chi) \leqslant \rho(\varphi, \psi) + \rho(\psi, \chi).$$

5. Sea G un conjunto abierto medible según Jordan del espacio euclídeo n-dimensional R^n. Un conjunto X de las funciones continuas en la clausura $\overline{G}$ del conjunto G forma un espacio métrico, si la distancia entre las funciones $\varphi \in X$ y $\psi \in X$ se determina mediante la fórmula

$$\rho(\varphi, \psi) = \int |\psi(x) - \varphi(x)|\, dG.$$

En efecto, si $\rho(\varphi, \psi) = 0$, es decir, si $\int |\psi(x) - \varphi(x)|\, dG = 0$, entonces, en virtud del corolario de la propiedad 9° de las integrales múltiples (véase el p. 44.6), $\varphi(x) = \psi(x)$ para todo $x \in G$ y, por ende, para todo $x \in \overline{G}$. La propiedad 2° de la distancia es, en este caso, evidente, y la propiedad 3° se comprueba con facilidad: si φ, ψ y χ son continuas en $\overline{G}$, se tiene

$$\rho(\varphi, \psi) = \int |\varphi(x) - \chi(x)|\, dG = \int |[\varphi(x) - \psi(x)] - [\psi(x) - \chi(x)]|\, dG \leqslant$$
$$\leqslant \int |\varphi(x) - \psi(x)|\, dG + \int |\psi(x) - \chi(x)|\, dG = \rho(\varphi, \psi) + \rho(\psi, \chi).$$

En el caso en que $n = 1$ y $\overline{G} = [a, b]$, la métrica introducida para las funciones continuas en el segmento $[a, b]$ tiene por expresión

$$\rho(\varphi, \psi) = \int_a^b |\varphi(x) - \psi(x)|\, dx. \tag{57.2}$$

Un espacio semejante también se introduce del modo natural para las funciones que están definidas en un intervalo infinito. Por ejemplo, cuando $a = -\infty$, $b = +\infty$, la distancia para dos funciones continuas y absolutamente integrables en todo el eje numérico, φ y ψ, se determina según la fórmula

$$\rho(\varphi, \psi) = \int_{-\infty}^{+\infty} |\varphi(x) - \psi(x)|\, dx. \tag{57.3}$$

Todo subconjunto de un espacio métrico X es, a su vez, un espacio métrico respecto de la misma métrica y se llama *subespacio del espacio* X.

Definición 2. *Dos espacios métricos X y X' se llaman isométricos, si entre dos puntos suyos existe una correspondencia biunívoca f que mantiene inalterable la distancia, es decir, una correspondencia tal que si*

$$x' = f(x), \quad y' = f(y), \quad x \in X, \quad y \in X, \quad x' \in X', \quad y' \in X',$$

entonces $\rho(x, y) = \rho(x', y')$ (tales correspondencias se denominan isométricas).

Definición 3. *Sea X un espacio métrico; la sucesión de sus puntos $\{x_n\}$ se denomina convergente al punto $x \in X$, si $\lim_{n \to \infty} \rho(x, x_n) = 0$, es decir, si para todo número $\varepsilon > 0$ existe un número n_ε tal que para cualesquiera números $n \geqslant n_\varepsilon$ se verifica la*

desigualdad $\rho(x, x_n) < \varepsilon$. *En este caso se escribe* $x = \lim\limits_{n \to \infty} x_n$, *o bien* $x_n \to x$ *cuando* $n \to \infty$, *y se dice que el punto x es el límite de la sucesión dada.*

Por ejemplo, la convergencia en los espacios métricos considerados en los casos 1 y 2 significa una convergencia ordinaria de las sucesiones numéricas (reales o complejas, respectivamente). En el ejemplo 3 la convergencia de la sucesión está representada por la convergencia de una sucesión de puntos en el espacio n-dimensional con la que ya nos hemos encontrado anteriormente (véase el p. 18.1). Es un espacio métrico de las funciones definidas y acotadas en cierto conjunto, donde la distancia entre las funciones se determina según la fórmula (57.1), la sucesión de funciones $\{\varphi_n\}$ converge hacia la función φ, si

$$\lim_{n \to \infty} \sup_{t \in X} |\varphi(t) - \varphi_n(t)| = 0,$$

es decir, si la sucesión de funciones $\{\varphi_n\}$ converge uniformemente en el conjunto X hacia la función φ (véase el t. 1; p. 36.2).

Por fin, el ejemplo 5 proporciona el tipo de convergencia de las funciones en el sentido de cierta métrica integral. Cuando $n = 1$, la convergencia es similar a la que se ha encontrado en el p. 55.2 (lema 2) y en el p. 56.7 (corolario del lema 4).

Ejercicio 1. El conjunto E de un espacio métrico X se llama acotado, si

$$d(E) \stackrel{\text{def}}{=} \sup_{x \in E,\, y \in E} \rho(x, y) < +\infty,$$

la magnitud $d(E)$ se denomina diámetro del conjunto E. Demuéstrese que toda sucesión convergente de un espacio métrico es acotada.

Definición 4. *La sucesión* $\{x_n\}$ *de puntos de un espacio métrico X se llama fundamental, si para cualquier número* $\varepsilon > 0$ *existe tal número* n_ε *que se cumple la desigualdad*

$$\rho(x_n, x_m) < \varepsilon,$$

cualesquiera que sean los números $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$.

Lema 1. *Si la sucesión* $\{x_n\}$ *converge, es fundamental.*

DEMOSTRACIÓN. Sea $\lim\limits_{n \to \infty} x_n = x$. En este caso para cualquier número $\varepsilon > 0$ existe tal número n_ε que para todos los números $n \geqslant n_\varepsilon$ se verifica la desigualdad

$$\rho(x, x_n) < \frac{\varepsilon}{2}.$$

Por consiguiente, si $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$, entonces

$$\rho(x_n, x_m) \leqslant \rho(x_n, x) + \rho(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon. \quad \square$$

Definición 5. *Un espacio métrico se llama completo, si toda sucesión fundamental de sus puntos converge hacia un punto que pertenece al mismo espacio.*

Es evidente que un espacio métrico que es isométrico respecto del espacio completo es también un espacio métrico completo.

Ejemplos. 6. Los espacios métricos de los números reales y complejos representan los ejemplos de espacios métricos completos. El espacio euclídeo n-dimen-

sional R^n (véase el p. 18.1) es también completo. Los números racionales ofrecen un ejemplo de espacio métrico incompleto.

7. Consideraremos un espacio métrico de las funciones definidas y acotadas en el conjunto X, donde la distancia entre las funciones se determina mediante la fórmula (57.1). En este espacio la sucesión de funciones φ_n, $n = 1, 2, \ldots$, es fundamental, siempre que para cualquier número $\varepsilon > 0$ existe un número n_ε tal que se verifica la desigualdad

$$\rho(\varphi_n, \varphi_m) = \sup_X |\varphi_n(x) - \varphi_m(x)| < \varepsilon$$

cualesquiera que sean los números $n \geqslant n_\varepsilon$ y $m \geqslant m_\varepsilon$, es decir, si la sucesión $\{\varphi_n\}$ satisface el criterio de Cauchy sobre la convergencia uniforme de la sucesión en el conjunto X (véase el p. 36.2). En vista de este criterio, la sucesión $\{\varphi_n\}$ converge uniformemente en el conjunto X hacia cierta función φ, es decir,

$$\lim_{n \to \infty} \sup_X |\varphi(x) - \varphi_n(x)| = 0. \tag{57.4}$$

Mostremos que esta función φ es también acotada y, por lo tanto, pertenece al espacio en consideración. En efecto, en virtud de (57.4), para todo número $\varepsilon > 0$, en particular para $\varepsilon = 1$, existe un número n_1 tal que se verifica la desigualdad

$$|\varphi(x) - \varphi_{n_1}(x)| < 1;$$

cualesquiera que sean $n \geqslant n_1$ y $x \in E$, razón por la cual

$$|\varphi(x)| \leqslant |\varphi(x) - \varphi_{n_1}(x)| + |\varphi_{n_1}(x)| < 1 + \sup_X |\varphi_{n_1}(x)|.$$

Puesto que la función φ_{n_1} está acotada, lo es también la función φ. Hemos demostrado de este modo que el espacio de funciones que se considera es completo.

Se puede probar que el espacio métrico de funciones consideradas en el ejemplo 5 no es completo.

Para cualquier espacio métrico X se introduce de modo natural el concepto de ε-entorno $U(x, \varepsilon)$ del punto $x \in X$, $\varepsilon > 0$:

$$U(x, \varepsilon) = \{y : y \in R, \quad \rho(y, x) < \varepsilon\},$$

y luego textualmente, al igual que para el espacio n-dimensional R^n (véase el t. 1, p. 18.2), se introducen los conceptos de punto adherente de un conjunto, de puntos límites y aislado, de puntos interior y de frontera, el concepto de clausura $\bar{A}$ del conjunto A, el de conjuntos cerrado y abierto.

Para los espacios métricos arbitrarios son válidos también los lemas 3, 4, 5 y 6, demostrados en el p. 18.2, para los conjuntos abiertos y cerrados de los espacios euclídeos n-dimensionales, con la particularidad de que las demostraciones aducidas en el p. 18.2, quedan en vigor también en el caso general.

Definición 6. *El conjunto A de un espacio métrico X se llama denso en el espacio X, si la clausura $\bar{A}$ del conjunto A coincide con el espacio X: $\bar{A} = X$.*

Por ejemplo, un conjunto de los números racionales es denso en el conjunto de números reales.

Es evidente que la propiedad de un conjunto de ser denso en el espacio se conserva cuando se realizan las aplicaciones isométricas de este espacio.

Definición 7. *El espacio métrico completo X^* se llama completación del espacio métrico X, si en el espacio X^* existe un subconjunto X', denso en X^* e isométrico respecto del espacio X.*

Por ejemplo, el conjunto de números reales es una completación del conjunto de números racionales.

A veces resulta cómodo "identificar" los elementos de los espacios X y X' que corresponden uno al otro en la correspondencia isométrica de los espacios X y X', y considerar, de este modo, el conjunto X como un subconjunto de su completación X^*. Explicaremos más detalladamente la operación de identificación de los elementos de dos espacios isométricos X e Y. Sean X e Y^* unos espacios métricos, $Y \subset Y^*$, y f: $X \to Y$, la aplicación isométrica. Examinemos el conjunto $X^* = X \cup (Y^* \setminus Y)$, que se obtiene del espacio X por adición a este último del conjunto $Y^* \setminus Y$. De este modo: $X^* \setminus X = Y^* \setminus Y$. Definamos para los puntos $x \in X^*$ e $y \in X^*$ el concepto de distancia $\rho_{X^*}(x, y)$. Introduzcamos, con el fin de comodidad, la aplicación F: $X^* \to Y^*$, que se da por la fórmula

$$F(x) \stackrel{\text{def}}{=} \begin{cases} f(x), & \text{si } x \in X, \\ x, & \text{si } x \in X^* \setminus X. \end{cases}$$

Está claro que F es una aplicación biunívoca (biyección) del conjunto X^* sobre Y^*.

Ahora, para cualesquiera $x \in X^*$ e $y \in Y^*$ pongamos

$$\rho_{X^*}(x, y) = \rho(F(x), F(y)).$$

Es fácil comprobar que la función $\rho_{X^*}(x, y)$, definida del modo indicado, satisface tres axiomas de distancia y, por lo tanto, X^* es un espacio métrico, mientras que la aplicación F aplica isométricamente el espacio X^* sobre Y^*, con la particularidad de que, realizándose dicha aplicación, el conjunto X se convierte en Y. Por eso, si el conjunto Y fue denso en el espacio Y^*, entonces el conjunto X será denso en el espacio X^*.

Cuando decimos "identifiquemos en el espacio Y^* el conjunto X con el espacio Y, que es isométrico respecto de X", sobreentendemos el estudio del espacio X^* en lugar de Y^*.

Mostremos que para cualquier espacio métrico incompleto existe su completación, es decir, probemos que todo espacio métrico incompleto es un subconjunto denso en cierto espacio métrico completo.

Teorema 1. *Para todo espacio métrico existe su completación.*

DEMOSTRACIÓN.

1. ESTRUCTURA DE LA COMPLETACIÓN X^* DEL ESPACIO MÉTRICO DADO X.

Dos sucesiones $\{x_n\}$ e $\{y_n\}$ de elementos del espacio X se llamarán *equivalentes*, si

$$\lim_{n \to \infty} \rho(x_n, y_n) = 0. \tag{57.5}$$

La equivalencia entre dos sucesiones $\{x_n\}$ e $\{y_n\}$ se designa mediante el símbolo $\{x_n\} \sim \{y_n\}$; la equivalencia posee las siguientes propiedades:

1°. Cualquier sucesión $\{x_n\}$ es equivalente a sí misma: $\{x_n\} \sim \{x_n\}$.

2°. Si $\{x_n\} \sim \{y_n\}$, entonces $\{y_n\} \sim \{x_n\}$.

3°. Si $\{x_n\} \sim \{y_n\}$ e $\{y_n\} \sim \{z_n\}$, entonces $\{x_n\} \sim \{z_n\}$.

Para nosotros serán de interés sólo las sucesiones fundamentales del espacio X. El conjunto de éstas se descompone en las clases disjuntas de las sucesiones equivalentes entre sí. Designemos estas clases por $x^*, y^*, z^*, \ldots$, y su totalidad mediante X^*. Si una sucesión fundamental $\{x_n\}$ está contenida en la clase x^*, esto se escribirá, como siempre, de la manera siguiente: $\{x_n\} \in x^*$.

II. DETERMINACIÓN DE LA DISTANCIA $\rho^*(x^*, y^*)$ EN X^*.

Sean $\{x_n\}$ e $\{y_n\}$ dos sucesiones fundamentales del espacio métrico X. La sucesión numérica $\rho(x_n, y_n)$ será también fundamental, es decir, satisfará la condición de Cauchy (véase el p. 4.7). En efecto, para cualesquiera números n y m

$$\rho(x_n, y_n) \leqslant \rho(x_n, x_m) + \rho(x_m, y_m) + \rho(y_m, y_n)$$

y, por lo tanto, en virtud de la simetría de los índices n y m,

$$|\rho(x_n, y_n) - \rho(x_m, y_m)| \leqslant \rho(x_n, x_m) + \rho(y_n, y_m). \tag{57.6}$$

De lo que las sucesiones $\{x_n\}$ e $\{y_n\}$ son fundamentales se deduce que para cualquier número $\varepsilon > 0$ existe un número n_ε tal que se cumplen las desigualdades

$$\rho(x_n, x_m) < \frac{\varepsilon}{2}, \quad \rho(y_n, y_m) < \frac{\varepsilon}{2} \tag{57.7}$$

cualesquiera que sean los números $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$. De (57.6) y (57.7), para $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$, obtenemos

$$|\rho(x_n, y_n) - \rho(x_m, y_m)| < \varepsilon.$$

Por consiguiente, la sucesión numérica $\{\rho(x_n, y_n)\}$ es fundamental, es decir, satisface la condición de Cauchy y, por ende, es convergente.

Sea $\{x_n\} \in x^*$, $\{y_n\} \in y^*$. Pongamos, por definición, $\rho^*(x^*, y^*) \stackrel{\text{def}}{=} \lim_{n \to \infty} \rho(x_n, y_n)$. En virtud de lo demostrado, el límite mencionado existe. Mostremos que la función $\rho^*(x^*, y^*)$, definida de esta forma, no depende de la elección de las sucesiones fundamentales $\{x_n\} \in x^*$ e $\{y_n\} \in y^*$ y satisface los axiomas de distancia.

Sea $\{x_n\} \in x^*$, $\{x'_n\} \in x^*$, $\{y_n\} \in y^*$, $\{y'_n\} \in y^*$. En este caso

$$\rho(x_n, y_n) \leqslant \rho(x_n, x'_n) + \rho(x'_n, y'_n) + \rho(y_n, y'_n)$$

y

$$|\rho(x_n, y_n) - \rho(x'_n, y'_n)| \leqslant \rho(x_n, x'_n) + \rho(y_n, y'_n).$$

Por ser equivalentes las sucesiones $\{x_n\}$, $\{x'_n\}$, y, respectivamente, $\{y_n\}$, $\{y'_n\}$, obtendremos (véase (57.5)):

$$\lim_{n \to \infty} \rho(x_n, x'_n) = \lim_{n \to \infty} \rho(y_n, y'_n) = 0$$

y, por consiguiente,

$$\lim_{n \to \infty} \rho(x_n, y_n) = \lim_{n \to \infty} \rho(x'_n, y'_n).$$

III. COMPROBACIÓN DE LOS AXIOMAS DE DISTANCIA PARA $\rho^*(x^*, y^*)$.

Sea $\{x_n\} \in x^*$, $\{y_n\} \in y^*$, $\{z_n\} \in z^*$. Si es que $\rho^*(x^*, y^*) = 0$, entonces $\lim_{n-\infty} \rho(x_n, y_n) = 0$, es decir, las sucesiones $\{x_n\}$ e $\{y_n\}$ son equivalentes lo que implica la coincidencia de los elementos x^* e y^*: $x^* = y^*$. De la igualdad $\rho(x_n, y_n) = \rho(y_n, x_n)$, pasando al límite para $n \to \infty$, obtendremos $\rho^*(x^*, y^*) = \rho^*(y^*, x^*)$, y de la desigualdad $\rho(x_n, y_n) \leqslant \rho(x_n, z_n) + \rho(z_n, y_n)$ obtenemos

$$\rho^*(x^*, y^*) \leqslant \rho^*(x^*, z^*) + \rho^*(z^*, y^*).$$

Así pues, X^* es un espacio métrico.

IV. CONSTRUCCIÓN DE UN SUBESPACIO DEL ESPACIO X^*, ISOMÉTRICO RESPECTO DEL ESPACIO X.

Sea $x \in X$. La sucesión $x_n = x$, $n = 1, 2, \ldots$, es obviamente fundamental. Pongamos en correspondencia a todo $x \in X$ un punto $x^* \in X^*$ tal que $\{x\} \in x^*$. Si, asumida la correspondencia citada, al punto x le corresponde el punto x^*, y al punto y, el punto y^*, entonces, evidentemente, para $x \neq y$ tendremos $x^* \neq y^*$, con la particularidad de que $\rho^*(x^*, y^*) = \lim_{n-\infty} \rho(x, y) = \rho(x, y)$, es decir, dicha correspondencia realiza una correspondencia isométrica biunívoca entre el espacio X y un cierto subconjunto X' del espacio X^*.

El punto x^* del espacio X^*, correspondiente al punto $x \in X$ en la correspondencia analizada lo denotaremos también, para simplificar, mediante x, y el espacio X', mediante X. Se puede considerar que los puntos correspondientes de los espacios X y X' quedan simplemente identificados (véase la observación que sigue la definición 7). En estas designaciones se tiene una inclusión isométrica

$$X \subset X^*.$$

V. DEMOSTRACIÓN DE LA DENSIDAD DE X EN X^*.

Probemos que todo punto x^* del espacio X^* es un punto de adherencia del conjunto X. Con este fin basta mostrar que para cualquier punto $x^* \in X^*$ existe una sucesión $x_n \in X$, $n = 1, 2, \ldots$, que converge hacia x^*.

Sean $x^* \in X^*$ y $\{x_n\} \in x^*$, $x_n \in X$. Un punto del espacio X^*, que contiene una sucesión fundamental cuyos términos son todos iguales a un mismo punto x_n, lo designaremos, de acuerdo con el convenio asumido anteriormente, también mediante x_n. Demostremos que la sucesión $\{x_n\}$, $x_n \in X^*$, converge al punto $x^* \in X^*$. Fijemos el número $\varepsilon > 0$. De lo que la sucesión $\{x_n\}$ es fundamental se deduce que existe un número n_ε tal que se cumple la desigualdad

$$\rho(x_m, x_n) < \frac{\varepsilon}{2} \tag{57.8}$$

cualesquiera que sean los números $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$. Observando que según la definición de distancia en X^*

$$\rho^*(x^*, x_n) = \lim_{m-\infty} \rho(x_m, x_n),$$

de la desigualdad (57.8) obtenemos, para $n \geqslant n_\varepsilon$,

$$\rho^*(x^*, x_n) \leqslant \frac{\varepsilon}{2} < \varepsilon,$$

es decir,

$$\lim_{n\to\infty} \rho^*(x^*, x_n) = 0,$$

lo que implica que x^* es un punto de adherencia del conjunto X. Así pues, $\bar{X} = X^*$.

VI. DEMOSTRACIÓN DE LA COMPLETITUD DEL ESPACIO X^*.

Sea $\{x_n^*\}$ una sucesión fundamental de puntos del espacio X^*, $x_n \in X$ y $\rho^*(x_n^*, x_n) < \frac{1}{n}$, $n = 1, 2, \ldots$. Tales puntos x_n existen en vista de que X es denso en X^*.

La sucesión $\{x_n\}$ es fundamental. Efectivamente, notando que

$$\rho^*(x_n, x_m) \leqslant \rho^*(x_n, x_n^*) + \rho^*(x_n^*, x_m^*) + \rho^*(x_m^*, x_m) < \frac{1}{n} + \rho^*(x_n^*, x_m^*) + \frac{1}{m},$$

elijamos el número n_ε de modo tal que sea

$$\rho^*(x_n^*, x_m^*) < \frac{\varepsilon}{3}, \quad \frac{1}{n} < \frac{\varepsilon}{3}, \quad \frac{1}{m} < \frac{\varepsilon}{3}$$

para cualesquiera $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$. Entonces

$$\rho(x_n, x_m) = \rho^*(x_n, x_m) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \tag{57.9}$$

para cualesquiera $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$, es decir, la sucesión $\{x_n\}$ es fundamental.

Denotemos mediante x^* la clase de sucesiones equivalentes a la que pertenece la sucesión $\{x_n\}$. Es obvio que

$$\rho^*(x^*, x_n^*) \leqslant \rho^*(x^*, x_n) + \rho^*(x_n, x_n^*) = \rho^*(x^*, x_n) + \frac{1}{n}.$$

Pero, de (57.9) obtenemos, cuando $m \to \infty$ y $n \geqslant n_\varepsilon$:

$$\rho^*(x^*, x_n) = \lim_{m\to\infty} \rho(x_m, x_n) \leqslant \varepsilon.$$

Por consiguiente,

$$\lim_{n\to\infty} \rho^*(x^*, x_n) = 0,$$

por lo cual también

$$\lim_{n\to\infty} \rho^*(x^*, x_n^*) = 0.$$

De este modo, hemos demostrado que la sucesión fundamental dada $\{x_n^*\}$ converge en X^*. La completitud de X^* queda demostrada. □

Ejercicio 2. Demuéstrese que la completación de un espacio métrico es única salvo unos espacios isométricos.

Definición 8. *Una función numérica f (de valores reales o complejos), definida en el conjunto A del espacio métrico X, se llama continua en el punto* $x_0 \in A$ (*o, más*

detalladamente, continua respecto del conjunto A en el punto $x_0 \in A$), si para todo número $\varepsilon > 0$ existe un número $\delta = \delta(\varepsilon)$ tal que se cumple la desigualdad

$$|f(x) - f(x_0)| < \varepsilon$$

cualesquiera que sean los puntos $x \in U(x_0, \delta) \cap A$.

Definición 9. *Una función f, definida en el conjunto A del espacio métrico X, se denomina continua en el conjunto $B \subset A$, si es continua respecto del conjunto A en todo punto $x_0 \in B$.*

Ejercicio 3. Enúnciese, con ayuda del concepto de sucesión, la definición de la continuidad en el punto x_0 de la función f, definida en el conjunto $A \ni x_0$, y demuéstrese que esta definición es equivalente a la definición 8.

Al igual que en el p. 36.4 (véase el t. 1), se demuestra textualmente que el límite de una sucesión de funciones continuas convergente uniformemente en el espacio métrico es una función continua.

Ejemplo. Consideremos un espacio métrico de funciones f, acotadas y continuas en cierto *espacio métrico X*, la distancia entre las cuales se determina según la fórmula (57.1). Por cuanto el carácter fundamental de la sucesión $[f_n]$ en el sentido de la métrica (57.1) significa que esta sucesión satisface la condición de Cauchy de convergencia uniforme en el conjunto X, entonces toda sucesión fundamental de las funciones continuas $[f_n]$ converge uniformemente hacia cierta función f. Esta función f es, según lo dicho anteriormente, continua y, de acuerdo con lo demostrado más arriba en este punto, acotada en X, es decir, pertenece al espacio de funciones que se considera.

De este modo, un *espacio de funciones acotadas y continuas en el espacio métrico X es espacio métrico completo*. Es, evidentemente, un subespacio de todas las funciones acotadas en X cuya distancia se ha determinado según la fórmula (57.1).

En particular, ya que toda función continua en cierto compacto A, dispuesto en el espacio euclídeo n-dimensional R^n, está acotada (véase el p. 19.4), entonces el espacio de funciones continuas en el compacto A con la distancia definida según la fórmula (57.1) es completo.

Definición 10. *Sea X un espacio métrico. La función f, definida en un conjunto de pares ordenados (x, y), donde $x \in A$, $y \in B$, $A \subset X$, $B \subset X$, se llama continua en el punto (x_0, y_0), $x_0 \in A$, $y_0 \in B$, si para cualquier número $\varepsilon > 0$ existe un número $\delta = \delta(\varepsilon) > 0$ tal que se verifica la desigualdad $|f(x, y) - f(x_0, y_0)| < \varepsilon$ cualesquiera que sean los pares (x, y) de tal género que $x \in U(x_0, \delta) \cap A$, $y \in U(y_0, \delta) \cap B$.*

Una función, continua en todo punto (x, y) de cierto conjunto de pares, se denomina continua en este conjunto.

Ejercicios. 4. Compruébense los axiomas de distancia para la función $\rho(\varphi, \psi)$ definida mediante la fórmula (57.3) para el espacio de funciones absolutamente integrables y continuas en todo el eje numérico.

5. Dése un ejemplo de sucesión de las funciones continuas que converge en cierto segmento en el sentido de la distancia (57.2), pero no converge en el mismo segmento en el sentido de convergencia en un punto (es decir, en el sentido de la definición 3, el p. 36.1).

6. Dése un ejemplo de sucesión que converge en cierto segmento en el sentido de convergencia en un punto, pero que no converge en el mismo segmento en el sentido de la distancia (57.2).

7. Demuéstrese que un espacio de las funciones continuas en el segmento $[a, b]$, la distancia entre las cuales se determina por la fórmula (57.2), no es completo.

57.2. ESPACIOS LINEALES

Definición 11. *Un conjunto* $X = \{x, y, z, \ldots\}$ *se llama espacio lineal real (o espacio vectorial sobre un campo de números reales), si:*

a todo par ordenado (x, y) *de elementos* $x \in X$ *e* $y \in X$ *se le ha puesto en correspondencia un elemento del espacio* X, *llamado suma de* x *e* y *y denotado por* $x + y$;

a todo elemento $x \in X$ *y a todo número real* λ *se les ha puesto en correspondencia el único elemento del espacio* X, *llamado producto de* λ *por* x *y denotado mediante* λx.

En este caso se cumplen los siguientes grupos de axiomas:

1. a) $x + y = y + x$ *para cualesquiera* $x \in X$ *e* $y \in X$;
   b) $x + (y + z) = (x + y) + z$ *para cualesquiera* $x \in X$, $y \in X$ *y* $z \in X$;
   c) *en* X *existe un elemento, llamado nulo y denotado por* 0, *tal que* $x + 0 = x$ *para cualquier* $x \in X$;
   d) *para todo* $x \in X$ *existe un elemento del conjunto* X, *llamado opuesto al elemento* x, *que se designa mediante* $-x$ *y es tal que* $x + (-x) = 0$.
2. a) $1x = x$ *para cualquier* $x \in X$;
   b) $\lambda(\mu x) = (\lambda\mu)\, x$ *para todo* $x \in X$ *y cualesquiera números reales* λ *y* μ.
3. a) $(\lambda + \mu)x = \lambda x + \mu x$ *para todo* $x \in X$ *y cualesquiera números reales* λ *y* μ;
   b) $\lambda(x + y) = \lambda x + \lambda y$ *para cualesquiera* $x \in X$, $y \in Y$ *y cualquier número real* λ.

Para todo par de elementos $x \in X$ *e* $y \in Y$ el elemento $x + (-y)$ se llama diferencia de los elementos x e y, y se denota mediante $x - y$.

Si en la definición aducida de espacio lineal real se sustituyen todos los números reales por los complejos: $\lambda, \mu \in C$; se obtendrá, entonces, la definición del *espacio lineal complejo.*

Ejemplos. 1. El conjunto de todos los números reales (complejos) forma un espacio lineal real (complejo).

2. Sea X cierto conjunto. La totalidad $F(X)$ de todas las funciones $f: X \to R$ ($f: X \to C$, respectivamente), para la definición natural de su sumación y multiplicación por los números reales (complejos):

$$(f_1 + f_2)(x) \stackrel{\text{def}}{=} f_1(x) + f_2(x), \quad (\lambda f)(x) \stackrel{\text{def}}{=} \lambda(f(x)),$$
$$f_1 \in F(X), \quad f_2 \in F(X), \quad f \in F(X), \quad \lambda \in R \quad \text{o bien} \quad \lambda \in C,$$

será un espacio lineal real (complejo).

3. El conjunto de todos los polinomios de una variable con coeficientes reales (complejos) es un espacio real (complejo) lineal.

4. El conjunto de todos los polinomios de grados, no superiores a n natural, de una variable con los coeficientes reales (complejos) es un espacio lineal real (complejo).

5. Un espacio de toda clase de sucesiones numéricas $\{x_n\}$, $x_n \in R$ (o bien $x_n \in C$), $n \in N$, para la definición natural de la operación de su sumación y multiplicación por un número (véase el p. 4.9) es también espacio lineal.

Definición 12. *Un conjunto X', contenido en el espacio lineal X (real o complejo), se llama subespacio de este espacio, si todas las combinaciones lineales de elementos del conjunto X' se contienen en él.*

En otras palabras, el conjunto $X' \subset X$ es un subespacio del espacio X, si para cualesquiera dos elementos $x \in X'$, $y \in X'$ y cualesquiera números $\lambda \in R$, $\mu \in R$ ($\lambda \in C$, $\mu \in C$, respectivamente) tiene lugar la inclusión

$$\lambda x + \mu y \in X'.$$

Es evidente que el subespacio X' del espacio lineal X es, a su vez, un espacio lineal.

Si X es un espacio lineal y $x \in X$, entonces la totalidad de todos los elementos del *espacio* X del tipo λx, donde λ son números cualesquiera, sirve de ejemplo de subespacio del espacio X.

El conjunto de las funciones de valores reales y continuas en cierto conjunto $X \subset R^n$ es un subespacio del espacio de todas las funciones de valores reales, definidas en X.

Los elementos de los espacios lineales se llaman, comúnmente, *puntos* o *vectores*.

Definición 13. *Un sistema finito de vectores $x_1, \ldots, x_n$ del espacio lineal X (real o complejo) se llama linealmente dependiente, si existen tales números $\lambda_1, \ldots, \lambda_n$ (reales o complejos, respectivamente), no todos iguales a cero, que*

$$\lambda_1 x_1 + \ldots + \lambda_n x_n = 0.$$

En el caso contrario, es decir, cuando de la igualdad citada se deduce que todos los números $\lambda_1, \lambda_2, \ldots, \lambda_n$ son nulos, el sistema de vectores $x_1, \ldots, x_n$ se denomina linealmente independiente.

Definición 14. *Un sistema de vectores x_α, $\alpha \in \mathfrak{A}$ ($\mathfrak{A}$ es un conjunto de índices), del espacio lineal X se llama linealmente independiente, si cualquiera de sus subsistemas finitos $x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha_n}$ es linealmente independiente.*

Ejercicios. 8. Demuéstrese que si el sistema x_α, $\alpha \in \mathfrak{A}$, es linealmente independiente, entonces $x_\alpha \neq 0$ para cualesquiera $\alpha \in \mathfrak{A}$.

9. Demuéstrese que para que un sistema finito de vectores sea linealmente dependiente, es necesario y suficiente que por lo menos uno de los vectores sea una combinación lineal de los demás.

Definición 15. *Sea dado un conjunto $\{x_\alpha, \alpha \in \mathfrak{A}\}$ de vectores del espacio lineal X. La totalidad de toda clase de combinaciones lineales finitas de los elementos de este conjunto, es decir, la totalidad de toda clase de vectores del tipo*

$$\lambda_1 x_{\alpha_1} + \lambda_2 x_{\alpha_2} + \ldots + \lambda_k x_{\alpha_k},$$

donde $x_{\alpha_j} \in \{x_\alpha, \alpha \in \mathfrak{A}\}$ y λ_j son los números, $j = 1, 2, \ldots, k$, lleva el nombre de cápsula lineal del conjunto $\{x_\alpha, \alpha \in \mathfrak{A}\}$.

Definición 16. *Si en el espacio X (real o complejo) se tiene un sistema de n vectores linealmente independientes cuya cápsula lineal está representada por el espacio X, entonces dicho espacio se denomina n-dimensional y se denota por R^n, mientras que todo sistema ordenado de n vectores linealmente independientes cuya cápsula lineal está representada por el espacio R^n, se llama base del espacio.*

En otras palabras, los vectores $e_1, e_2, \ldots, e_n$ son la base del espacio R^n, si:

1) los vectores $e_1, e_2, \ldots, e_n$ son linealmente independientes;

2) para todo $x \in R^n$ existen tales números $\lambda_1, \lambda_2, \ldots \lambda_n$ que $x = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n$.

Los elementos del espacio R^n se denominan vectores n-dimensionales (reales o complejos, respectivamente).

Cualquier espacio n-dimensional se llama espacio de dimensión finita.

Ejercicios. 10. Demuéstrese que en un espacio n-dimensional cualquier sistema de vectores linealmente independientes, cuya cápsula lineal está representada por todo el espacio, se compone de n vectores.

11. Demuéstrese que todo sistema de n vectores linealmente independientes en un espacio n-dimensional es la base de éste.

Un ejemplo de espacio real n-dimensional lo constituye el espacio vectorial aritmético n-dimensional (véase el p. 18.4).

Por analogía con este último espacio puede construirse un espacio aritmético complejo n-dimensional C^n. Se denominan puntos de este espacio los sistemas ordenados de n números complejos: $x = (x_1, \ldots, x_n)$ $x_j \in C$, $j = 1, 2, \ldots, n$. En este caso, si $x \in C^n$, $\lambda \in C$, se tiene

$$\lambda x \stackrel{\text{def}}{=} (\lambda x_1, \ldots, \lambda x_n),$$

y para $x = (x_1, \ldots, x_n) \in C^n$ e $y = (y_1, \ldots, y_n) \in C^n$

$$x + y \stackrel{\text{def}}{=} (x_1 + y_1, \ldots, x_n + y_n).$$

La base en este espacio la constituyen los vectores $e_i = (\delta_1^i, \ldots, \delta_n^i)$, donde δ_j^i es el así llamado *símbolo de Kronecker* *)

$$\delta_j^i = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$$

Es evidente que $x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i$.

Como otro ejemplo de espacio lineal de dimensión finita interviene el espacio $\mathscr{P}^n$ de los polinomios de grados no superiores a n natural. Es $(n + 1)$-dimensional: su dimensión es igual al número de coeficientes que tienen los polinomios en consideración.

Definición 17. *La aplicación f de un espacio lineal X en otro espacio lineal Y se denomina aplicación lineal (o, que es lo mismo, operador lineal), si para cuales-*

*) L. Kronecker (1823 — 1891), matemático alemán.

quiera dos elementos y cualesquiera números λ *y* μ *se verifica la igualdad*

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

El conjunto de todos los operadores lineales $f: X \to Y$, que aplican el espacio lineal X en el espacio lineal Y, se designará mediante $\mathscr{L}(X, Y)$. Por comprobación inmediata nos convencemos con facilidad de que el conjunto $\mathscr{L}(X, Y)$, para la definición natural de sumación de sus elementos y multiplicación de ellos por un número, es decir, para la definición de estas operaciones según las fórmulas

$$(f_1 + f_2)(x) \stackrel{\text{def}}{=} f_1(x) + f_2(x), \quad f_1 \in \mathscr{L}(X, Y), \quad f_2 \in \mathscr{L}(X, Y),$$
$$(\lambda f)(x) \stackrel{\text{def}}{=} \lambda(f(x)), \quad f \in \mathscr{L}(X, Y), \quad \lambda \in R \text{ o } \lambda \in C$$
$$x \in X,$$

forma también un espacio lineal (real, si los espacios X e Y eran espacios lineales reales, o complejo, si X e Y eran complejos).

Definición 18. *Si* $f: X \to Y$ *e* Y *es un espacio lineal, entonces el conjunto* $\{x: f(x) = 0\} \subset X$ *se llama núcleo de la aplicación* f *y se denota mediante* $\ker f$ **):*

$$\ker f \stackrel{\text{def}}{=} \{x: f(x) = 0\}.$$

Lema 2. *Para que la aplicación lineal* $f: X \to Y$ *de un espacio lineal* X *en otro espacio lineal* Y *sea una aplicación biunívoca de* X *en* Y*, es decir, sea una inyección, es necesario y suficiente que su núcleo se componga sólo en un elemento nulo:*

$$\ker f = 0.$$

DEMOSTRACIÓN DE LA NECESIDAD. Evidentemente, cualquier operador lineal f hace pasar cero en cero, pues para todo $x \in X$ tenemos: $f(0) = f(0x) = 0f(x) = 0$. Por eso, si f es una inyección, no existe $x \neq 0$ tal que sea $f(x) = 0$. Esto precisamente es un testimonio de que $\ker f = 0$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Sea $\ker f = 0$ y $f(x) = f(y)$. Entonces, por ser lineal la aplicación f, se tiene $f(x - y) = f(x) - f(y) = 0$, es decir, $x - y \in \ker f$ y, como $\ker f = 0$, se tiene $x - y = 0$. Por consiguiente $x = y$. Esto significa que f es una inyección. □

Como ejemplos de aplicaciones lineales biunívocas sirven las transformaciones directa e inversa de Fourier en los correspondientes espacios lineales de funciones (véanse los lemas 2 y 3 en el p. 56.5).

Definición 19. *Sean* X *e* Y *unos espacios lineales. La aplicación lineal biunívoca del espacio* X *sobre el espacio* Y *se denomina aplicación isomorfa o isomorfismo de los espacios lineales.*

Si para los espacios lineales X *e* Y *existe una aplicación isomorfa de* X *sobre* Y*, se denominan isomorfos.*

Dos espacios isomorfos pueden diferenciarse sólo en la naturaleza de sus elementos y no en las propiedades del espacio lineal en sí, por eso en lo que sigue los espacios lineales isomorfos no se distinguirán.

---

*) De la palabra inglés kernel (núcleo).

Ejercicio 12. Demuéstrese que todos los espacios lineales n-dimensionales son isomorfos entre sí.

Definición 20. *Un espacio lineal que no es de dimensión finita se llama espacio de dimensión infinita.*

Es obvio que un espacio lineal es de dimensión infinita cuando, y sólo cuando, no tiene una base finita.

Un ejemplo del espacio de dimensión infinita lo representa un espacio lineal de todos los polinomios de una sola variable. En efecto, este espacio está privado, a ciencia cierta, de base finita: cualquier combinación lineal del sistema finito dado de polinomios es un polinomio de grado no superior al del polinomio mayor del sistema citado, a consecuencia de lo cual los polinomios de grados superiores no pueden obtenerse por el procedimiento mencionado.

La tentativa de generalizar el concepto de base en el caso de los espacios de dimensión infinita lleva a sumas infinitas, es decir, a las series de la forma $\sum_{n=1}^{\infty} \lambda_n e_n$. Para que haya sentido de hablar de su suma en el espacio X, en éste ha de ser definido el concepto de convergencia de las sucesiones. El punto que sigue está dedicado a la consideración de uno de los espacios de este género.

57.3. ESPACIOS NORMALIZADOS Y SEMINORMALIZADOS

Definición 21. *Un espacio lineal X (real o complejo) se denomina normalizado, si en el conjunto de sus puntos está definida una función real, llamada norma y denotada con $\|x\|_X$, o, en la forma más breve, $\|x\|$, $x \in X$, y que posee las siguientes propiedades*:

1°) $\|x\| \geqslant 0$, $x \in X$;

2°) $\|\lambda x\| = |\lambda| \, \|x\|$, $x \in X$, λ *es un número*;

3°) $\|x + y\| \leqslant \|x\| + \|y\|$, $x \in X$, $y \in X$;

4°) *si* $\|x\| = 0$, *se tiene* $x = 0$.

Indiquemos que de la propiedad 2° se deduce que si $x = 0$, entonces $\|x\| = 0$. En efecto, fijando un elemento $x \in X$ arbitrario, obtendremos

$$\|0\| = \|0 \cdot x\| = 0 \|x\| = 0.$$

Definición 22. *Si en el conjunto de puntos de un espacio lineal X está definida una función real $\|x\|$, $x \in X$, que satisface sólo las propiedades 1°, 2°, 3°, entonces el espacio X se llama seminormalizado y la función $\|x\|$, seminorma.*

La propiedad 2° de la norma (seminorma) se llama *homogeneidad* de la norma (seminorma) y la propiedad 3°, *desigualdad triangular*.

Demos a conocer que todo subconjunto de un espacio lineal seminormalizado (en particular, normalizado) que representa un subespacio del espacio lineal es, a su vez, un espacio lineal seminormalizado (normalizado, respectivamente).

Ejercicio 13. Aclárese si serán ¿norma? ¿seminorma? ¿para qué funciones? ¿para qué n?

las expresiones $\sup_{a \leqslant t \leqslant b} |f^{(n)}(t)|$, $\int_a^b |f^{(n)}(t)| \, dt$.

57.4. EJEMPLOS DE ESPACIOS NORMALIZADOS Y SEMINORMALIZADOS

1. El conjunto de números reales y el de números complejos, si se toma por norma en ellos el valor absoluto de los números, forman espacios lineales normalizados.

2. Si en un espacio n-dimensional aritmético real R^n la norma del vector $x = (x_1, \ldots, x_n) \in R^n$ se define como su longitud (véase el p. 18.4)

$$\|x\| \stackrel{\text{def}}{=} |x| = \sqrt{x_1^2 + \ldots + x_n^2},$$

entonces R^n será un espacio lineal normalizado.

3. Un espacio n-dimensional aritmético complejo C^n (véase el p. 57.2) será normalizado, si ponemos

$$\|x\| \stackrel{\text{def}}{=} \sqrt{|x_1|^2 + \ldots + |x_n|^2}, \quad x = (x_1, \ldots, x_n) \in C^n.$$

4. En un espacio n-dimensional aritmético real R^n puede introducirse no sólo la norma coincidente con la longitud $|x|$ de sus elementos $x = (x_1, \ldots, x_n) \in R^n$. Por ejemplo, pongamos

$$\|x\|_p \stackrel{\text{def}}{=} (|x_1|^p + \ldots + |x_n|^p)^{1/p}, \quad 1 \leqslant p < +\infty,$$

$$\|x\|_\infty = \max_{i = 1, 2, \ldots, n} |x_i|.$$

Es evidente que la longitud del vector coincide con la norma $\|x\|_2$. Comprobemos el cumplimiento de los axiomas de las normas para $\|x\|_r$, $1 \leqslant r \leqslant +\infty$. Para $r = 1$, según la propiedad del valor absoluto de números se tiene

$$\|x + y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leqslant \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1.$$

Apliquemos la desigualdad de Minkowski para el caso en que $1 < p < +\infty$ (véase el p. 35.8*):

$$\|x + y\|_p = \left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{1/p} \leqslant \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p \right)^{1/p} =$$

$$= \|x\|_p + \|y\|_p.$$

Para $\|x\|_\infty$ tenemos

$$\|x + y\|_\infty = \max_{i = 1, 2, \ldots, n} |x_i + y_i| \leqslant \max_{i = 1, 2, \ldots, n} (|x_i| + |y_i|) \leqslant$$

$$\leqslant \max_{i = 1, 2, \ldots, n} |x_i| + \max_{i = 1, 2, \ldots, n} |y_i| = \|x\|_\infty + \|y\|_\infty.$$

Las demás propiedades de las normas para $\|x\|_r$, $1 \leqslant r \leqslant +\infty$, se comprueban de una manera más fácil.

Ejercicio 14. Demuéstrese que $\|x\|_\infty = \lim_{p \to \infty} \|x\|_p$, $x \in R^n$.

Definición 23. *Dos normas* $\|x\|$ *y* $\|x\|^*$ *en un espacio lineal normalizado* X *se llaman equivalentes, si existen unas constantes* $c_1 > 0$ *y* $c_2 > 0$ *tales, que para cualesquiera* $x \in X$ *se verifica la desigualdad*

$$c_1\|x\| \leqslant \|x\|^* \leqslant c_2\|x\|.$$

Teorema 2. *En un espacio lineal de dimensión finita todas las normas son equivalentes.*

DEMOSTRACIÓN. Sea X un espacio lineal de dimensión finita. Por consiguiente, existe en él una base $\{e_1, \ldots, e_n\}$ compuesta de cierto número $n \in N$ de sus elementos y para cualquier $x \in X$ se tiene una, y sólo una, base

$$x = x_1e_1 + \ldots + x_ne_n.$$

Sea $\|x\|$ una norma en el espacio X. Mostremos que es equivalente a la norma cuadrática

$$\|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Por cuanto dos normas, cada una de las cuales es equivalente a la tercera, son también equivalentes entre sí, de esto se deducirá que todas las normas de cualquier espacio de dimensión finita son equivalentes.

Hemos de notar, ante todo, que $c_1 \stackrel{\text{def}}{=} \|e_1\| + \ldots + \|e_n\| > 0$, pues para todo $k = 1, 2, \ldots, n$, se verifica la desigualdad $e_k \neq 0$, y, por ende, $\|e_k\| > 0$. Luego, de la desigualdad evidente

$$|x_k| \leqslant \sqrt{x_1^2 + \ldots + x_n^2} = \|x\|_2, \quad k = 1, 2, \ldots, n,$$

obtenemos, sirviéndose de la propiedad de la norma, la desigualdad

$$\|x\| = \|x_1e_1 + \ldots + x_ne_n\| \leqslant |x_1|\,\|e_1\| + \ldots + |x_n|\,\|e_n\| \leqslant$$
$$\leqslant (\|e_1\| + \ldots + \|e_n\|)\|x\|_2 = c_1\|x\|_2.$$

Así pues, existe tal $c_1 > 0$ que para todo $x \in X$ se tiene

$$\|x\| \leqslant c_1\|x\|_2.$$

Demostremos ahora que existe tal $c_2 > 0$ que

$$\|x\| \geqslant c_2\|x\|_2.$$

Por cuanto, siendo $x = 0$, esta desigualdad se cumple, evidentemente, para cualquier $c_2 > 0$, será suficiente demostrarla sólo para el caso en que $x \neq 0$. Elijamos una base $\{e_1, \ldots, e_n\}$ en el espacio X de manera tal que se componga de los vectores unidad en el sentido de la norma cuadrática

$$\|e_1\|_2 = \ldots = \|e_n\|_2 = 1.$$

Esto es siempre factible, puesto que si $\{e_1, \ldots, e_n\}$ es una base del espacio lineal y $\|\cdot\|$ es una norma en dicho espacio, entonces

$$\left\{\frac{e_1}{\|e_1\|}, \ldots, \frac{e_n}{\|e_n\|}\right\}$$

será también su base, con la particularidad de que la norma de todos los elementos suyos será igual a 1:

$$\left\| \frac{e_k}{\|e_k\|} \right\| = \frac{1}{\|e_k\|} \|e_k\| = 1, \quad k = 1, 2, \ldots, n.$$

El espacio X provisto de la base elegida puede considerarse como un espacio n-dimensional aritmético (véase el p. 18.4). Para esto es suficiente asignar a todo su vector $x = x_1 e_1 + \ldots + x_n e_n$ un surtido ordenado de n numeros $(x_1, \ldots, x_n)$ que representan sus coordenadas respecto de la base citada. En este caso la norma cuadrática $\|x\|_2$ será la longitud del vector x:

$$\|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2} = |x|.$$

La esfera unitaria $S^{n-1} = \{x: x_1^2 + \ldots + x_n^2 = 1\}$ de este espacio es, como se sabe (véase el p. 18.3 y 18.4), un compacto. Consideraremos en dicha esfera una función

$$f(x) \stackrel{\text{def}}{=} \|x\|.$$

De la desigualdad

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leqslant \|x - y\| \text{ *)} \leqslant$$
$$\leqslant c_1 \|x - y\|_2 = c_1 |x - y|, \quad x \in X, \quad y \in X,$$

se infiere que dicha función es continua en todo el espacio X y, por lo tanto, en la esfera S^{n-1}.

Por cuanto para cualquier punto $x \in S^{n-1}$ tenemos $\|x\|_2 = 1$, entonces $x \neq 0$, por lo cual, en virtud de la propiedad 4° de la norma, la función f satisface en la esfera S^{n-1} la desigualdad $f(x) = \|x\| > 0$. De acuerdo con el teorema de Weierstrass, toda función continua en un compacto alcanza en éste su valor mínimo. Supongamos que la función f admite su mínimo en la esfera S^{n-1} en el punto $x_0 \in S^{n-1}$. Pongamos

$$c_2 \stackrel{\text{def}}{=} \min_{x \in S^{n-1}} f(x) = f(x_0) > 0.$$

En este caso para cualquier $x \in S^{n-1}$ tendremos:

$$\|x\| = f(x) \geqslant f(x_0) = c_2.$$

Ahora, al observar que para todo $x \in X$, $x \neq 0$, el punto $\dfrac{x}{\|x\|_2}$ se dispone en la esfera S^{n-1}:

$$\left\| \frac{x}{\|x\|_2} \right\|_2 = \frac{1}{\|x\|_2} \|x\|_2 = 1$$

---

*) Hemos aprovechado aquí la desigualdad $\|x\| - \|y\| \leqslant \|x - y\|$. Es válida para cualesquiera elementos de un espacio seminormalizado y se deduce fácilmente de la propiedad 3° de la seminorma en la definición 21 (véase más abajo el lema 4 en el p. 57.5).

y, por lo tanto, para él se verifica $\left\| \frac{x}{\|x\|_2} \right\| \geqslant c_2$, obtendremos

$$\|x\| = \left\| \|x\|_2 \frac{x}{\|x\|_2} \right\| = \left\| \frac{x}{\|x\|_2} \right\| \|x\|_2 \geqslant c_2 \|x\|_2,$$

es decir,

$$\|x\| \geqslant c_2 \|x\|_2, \quad x \in X, x \neq 0.$$

La equivalencia de las normas $\|x\|$ y $\|x\|_2$ queda demostrada. □

5. Sea nuevamente $1 \leqslant p < +\infty$. Consideremos un subespacio lineal de todas las sucesiones $x = (x_1, \ldots, x_n, \ldots)$, $x_n \in R$ (o bien $x_n \in C$), compuesto de tales sucesiones, para las cuales

$$\|x\|_p \stackrel{\text{def}}{=} \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < +\infty. \tag{57.10}$$

La función $\|x\|_p$ es una norma, lo que se comprueba por analogía con el caso finito (véase el ejemplo 4), puesto que, en particular, la desigualdad de Minkowski es lícita también para las sumas infinitas.

Cuando todos los elementos de las sucesiones en consideración son números reales, su espacio con la norma (57.10) se denota con l_p.

6. En el p. 41.6 para el operador lineal $A: R^n \to R^m$ se ha introducido la norma según la fórmula (véase 41.41))

$$\|A\| = \sup_{|x| \leqslant 1} |Ax|, \quad x \in R^n.$$

Es realmente una norma, en el sentido de la definición del p. 57.3, en el espacio lineal $\mathcal{A}(R^n, R^m)$, lo que se deducirá de los razonamientos a seguir.

Supongamos que X e Y son espacios normalizados lineales arbitrarios y $A: X \to Y$ es un operador lineal. Pongamos

$$\|A\| \stackrel{\text{def}}{=} \sup_{\|x\| \leqslant 1} \|Ax\|, \tag{57.11}$$

donde $\|x\| = \|x\|_X$ y $\|Ax\| = \|Ax\|_Y$.

Cuando los espacios lineales X e Y se eligen arbitrariamente, puede suceder que la cota superior $\|A\|$, determinada por la igualdad (57.11), no será finita para todo operador lineal $A: X \to Y$.

Sea $\mathcal{A}(X, Y)$, como siempre (véase el p. 57.2), un conjunto de todos los operadores lineales A que aplican el espacio X en el Y y sea $\mathcal{A}_c(X, Y)$ un conjunto de aquellos de los operadores citados, para los cuales $\|A\| < +\infty$. Probemos que $\mathcal{A}_c(X, Y)$ es también un espacio lineal y $\|A\|$, la norma en él. Si $A \in \mathcal{A}_c(X, Y)$ y $B \in \mathcal{A}_c(X, Y)$, entonces

$$\|A + B\| = \sup_{\|x\| \leqslant 1} \|(A + B)x\| = \sup_{\|x\| \leqslant 1} \|Ax + Bx\| \leqslant$$

$$\leqslant \sup_{\|x\| \leqslant 1} (\|Ax\| + \|Bx\|) \leqslant \sup_{\|x\| \leqslant 1} \|Ax\| + \sup_{\|x\| \leqslant 1} \|Bx\| = \|A\| + \|B\| < +\infty,$$

y, por lo tanto, $A + B \in \mathscr{L}_c(X, Y)$. Para cualquier $\lambda \in R$ (o bien $\lambda \in C$, en el caso de espacios complejos)

$$\|\lambda A\| = \sup_{\|x\| \leqslant 1} \|\lambda A x\| = \sup_{\|x\| \leqslant 1} |\lambda| \, \|Ax\| =$$

$$= |\lambda| \sup_{\|x\| \leqslant 1} \|Ax\| = |\lambda| \, \|A\| < +\infty$$

y, por consiguiente, $\lambda A \in \mathscr{L}_c(X, Y)$. De este modo $\mathscr{L}_c(X, Y)$ de hecho es un espacio lineal.

Luego, es evidente que de (57.11) proviene inmediatamente que $\|A\| > 0$. Además, si $\|A\| = 0$, es decir, si $\sup_{\|x\| \leqslant 1} \|Ax\| = 0$, entonces para todo x tal que $\|x\| < 1$ tiene lugar la igualdad $\|Ax\| = 0$, y, por lo tanto, también $Ax = 0$. Mas, en este caso, en general, para todo $x \in X$ tenemos también $Ax = 0$. En efecto, si x es un elemento del espacio X tal que $\|x\| > 1$, entonces, a ciencia cierta, $x \neq 0$, quiere decir,

$$\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|} \|x\| = 1.$$

Por eso, en virtud de lo demostrado, $A\left(\frac{x}{\|x\|}\right) = 0$. De aquí $\frac{1}{\|x\|} Ax = 0$, y, por ende, $Ax = 0$, cualquiera que sea $x \in X$. Esto significa que $A = 0$. Así pues, $\|A\|$ es realmente una norma en el espacio $\mathscr{L}_c(X, Y)$.

Si el valor de $\|A\|$, definido por la fórmula (57.11), es infinito: $\|A\| = +\infty$, diremos que la norma del operador A es infinita.

La norma $\|A\|$ (tanto finita como infinita) puede obtenerse también por un método un tanto diferente. A saber, resulta que

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}, \quad x \in X. \tag{57.12}$$

Para demostrar esta fórmula indiquemos que

$$\sup_{\|x\| \leqslant 1} \|Ax\| = \sup_{\|y\| = 1} \|Ay\|. \tag{57.13}$$

Efectivamente, por una parte es evidente que

$$\sup_{\|x\| \leqslant 1} \|Ax\| \geqslant \sup_{\|y\| = 1} \|Ay\|,$$

pues, al aumentar un conjunto numérico, su cota superior sólo puede crecer. Por otra parte, para cualquier elemento $x \in X$, tal que $0 < \|x\| \leqslant 1$, pongamos $y \stackrel{\text{def}}{=} \frac{x}{\|x\|}$; entonces $\|y\| = 1$ y $\|Ay\| = \frac{1}{\|x\|} \|Ax\| \geqslant \|Ax\|$. De aquí

$$\sup_{\|x\| \leqslant 1} \|Ax\| \leqslant \sup_{\|y\| = 1} \|Ay\|.$$

De las desigualdades obtenidas se deduce precisamente la igualdad (57.13).

Ahora tenemos:

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \neq 0} \left\| A \frac{x}{\|x\|} \right\| = \sup_{\|y\| = 1} \|Ay\| = \sup_{\|y\| \leqslant 1} \|Ay\| = \|A\|,$$

es decir, la fórmula (57.12) queda también demostrada. De esta fórmula se deduce con toda la evidencia que para cualquier $x \in X$, $x \neq 0$,

$$\frac{\|A(x)\|}{\|x\|} \leqslant \|A\|,$$

y, por consiguiente, para cualquier $x \in X$ tiene lugar la desigualdad.

$$\|Ax\| \leqslant \|A\| \|x\|,$$

donde $\|x\|$ es la norma en el espacio X, $\|Ax\|$, la norma en el espacio Y, y $\|A\|$, la norma en el espacio $\mathcal{A}(X, Y)$. Esta igualdad es, obviamente, una generalización de la desigualdad (41.42) del p. 41.6.

Existe un enfoque más al concepto de norma de un operador, relacionado con el concepto de los así llamados operadores acotados.

Definición 24. *Un operador $A: X \to Y$ se denomina acotado, si existe una constante $c > 0$ tal que para todos los elementos $x \in X$ se verifica la desigualdad.*

$$\|Ax\| \leqslant c\|x\|.$$

Si A es un operador lineal acotado, todas las constantes $c > 0$ que poseen la propiedad citada están acotadas inferiormente por cero, por lo cual su conjunto cuenta con una cota inferior finita no negativa. Designémosla mediante c_0:

$$c_0 = \inf\{c: \|Ax\| \leqslant c\|x\|, x \in X\}.$$

Mostremos que

$$c_0 = \|A\|.$$

Ante todo indiquemos que para cualquier elemento $x \in X$ la desigualdad

$$\|Ax\| \leqslant c_0\|x\|$$

es lícita. Efectivamente, si se encontrara un elemento $x_0 \in X$ tal que sea $\|Ax_0\| > c_0\|x_0\|$, existiría cierto $\varepsilon > 0$, para el cual se verificaría la desigualdad $\|Ax_0\| > (c_0 + \varepsilon)\|x_0\|$. Sin embargo, esto no es posible, puesto que con arreglo a la definición de cota interior existe un número $c > 0$ tal que $c < c_0 + \varepsilon$ y para todo $x \in X$ se cumple la desigualdad $\|Ax\| < c\|x\|$. En particular, $\|Ax_0\| \leqslant c\|x_0\| < (c_0 + \varepsilon)\|x_0\|$. De este modo, la cota inferior c_0 satisface también la desigualdad, mediante la cual se determina la acotación del operador A. Por ello, en la definición de la constante c_0 la cota inferior puede ser sustituida por un mínimo

$$c_0 = \min\{c: \|Ax\| \leqslant c\|x\|, x \in X\}.$$

De la desigualdad $\|Ax\| \leqslant c_0\|x\|$, para $x \neq 0$, tenemos

$$\frac{\|Ax\|}{\|x\|} \leqslant c_0,$$

de donde

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \leqslant c_0, \quad x \in X.$$

El caso de la desigualdad estricta

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} < c_0, \quad x \in X,$$

no es posible, pues en tal caso existiría un número $\varepsilon > 0$ tal que

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} < c_0 - \varepsilon$$

y, por consiguiente, para todo $x \in X$, $x \neq 0$, sería válida con mayor razón la desigualdad

$$\frac{\|Ax\|}{\|x\|} < c_0 - \varepsilon, \quad \text{o bien} \quad \|Ax\| < (c_0 - \varepsilon)\|x\|, \quad x \in X,$$

lo que contradiría la elección de c_0 en calidad de una constante mínima que posee la propiedad $\|Ax\| \leqslant c\|x\|$, $x \in X$.

Así pues,

$$c_0 = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|.$$

Hablando metafóricamente, esta igualdad significa que el operador A está acotado cuando, y sólo cuando, tiene norma finita. De este modo, el conjunto de operadores acotados constituye el espacio $\mathscr{L}_c(X, Y)$.

En el p. 41.6 fue mostrado que todo operador lineal $A: X \to Y$ tiene una norma finita en el caso cuando los espacios normalizados lineales X e Y sean de dimensión finita y a título de normas en dichos espacios se hayan tomado normas cuadráticas $\|x\|_2$ e $\|y\|_2$, $x \in X$, $y \in Y$. Por cuanto en los espacios lineales de dimensión finita todas las normas son equivalentes (véase el teorema 2 en el ejemplo 4), de aquí se deduce que

todo operador lineal A, que aplica el espacio lineal de dimensión finita A en otro espacio lineal, también de dimensión finita Y, está acotado, cualesquiera que sea la elección de las normas en estos espacios, es decir, en este caso

$$\mathscr{L}(X, Y) = \mathscr{L}_c(X, Y).$$

7. Un espacio lineal de todas las funciones reales acotadas, definidas en el conjunto arbitrario X, que representa un subespacio del espacio $F(X)$ de todas las funciones reales $f: X \to R$ (véase el p. 57.2) se convierte en un espacio normalizado, siempre que se introduce en él una norma según la siguiente fórmula

$$\|f\|_\infty \stackrel{\text{def}}{=} \sup_{t \in X} |f(t)|. \tag{57.14}$$

Designemos este espacio mediante $S(X)$. En el caso en que X sea un espacio métrico, el subespacio del espacio $S(X)$, compuesto de las funciones f continuas en X, se

designará con $C(X)$*), y la norma (57.14) en dicho espacio se denotará, asimismo, mediante $\|f\|_C$.

Si X es un compacto en R^n, entonces (véase el teorema 3 en el p. 19.6)

$$\|f\|_C = \sup_{t \in X} |f(t)| = \max_{t \in X} |f(t)|.$$

En particular, esto es cierto para el espacio $C[a, b]$ de funciones continuas en el segmento $[a, b]$ de la recta numérica.

8. Sea fijado un número p, $1 \leqslant p < +\infty$. Estudiemos un conjunto de funciones f definidas en cierto segmento $[a, b]$ y tales que la integral

$$\int_a^b |f(x)|^p dx$$

converge. Es fácil comprobar que este conjunto forma un espacio lineal que se denota mediante $RL_p[a, b]$**).

Pongamos

$$\|f\|_p \stackrel{\text{def}}{=} \left[\int_a^b |f(t)|^p dt\right]^{1/p}. \tag{57.15}$$

Mostremos que (57.15) es una seminorma en $RL_p[a, b]$. Es evidente que de la fórmula (57.15) se deduce inmediatamente que $\|f\|_p \geqslant 0$. Con ello, de la condición $\|f\|_p = 0$ no se desprende que $f = 0$. En efecto, consideremos por ejemplo una función

$$f(x) = \begin{cases} 1 & \text{para } x = a, \\ 0 & \text{para } x \in (a, b]. \end{cases}$$

Está claro que $\|f\|_p = 0$, pero la función f no es idénticamente nula en el segmento $[a, b]$, razón por la cual no es cero del espacio lineal $RL_p[a, b]$.

Comprobemos la homogeneidad de la expresión (57.14): para cualesquiera $f \in RL_p[a, b]$ y cualquier $\lambda \in R$ (o bien $\lambda \in C$) tenemos

$$\|\lambda f\|_p = \left[\int_a^b |\lambda f(t)|^p dt\right]^{1/p} = |\lambda| \left[\int_a^b |f(t)|^p dt\right]^{1/p} = |\lambda| \, \|f\|_p.$$

Demostremos para (57.15) la desigualdad triangular. Cualesquiera que sean las funciones $f \in RL_p[a, b]$ y $g \in RL_p[a, b]$, de acuerdo con la desigualdad de Minkowski para las integrales (véase el p. 28.4*), obtendremos:

$$\|f + g\|_p = \left[\int_a^b |f(t) + g(t)|^p dt\right]^{1/p} \leqslant$$

$$\leqslant \left[\int_a^b |f(t)|^p dt\right]^{1/p} + \left[\int_a^b |g(t)|^p dt\right]^{1/p} = \|f\|_p + \|g\|_p.$$

---

*) C es la primera letra de la palabra latina "continuum".

**) R es la primera letra del apellido B. Riemann; L, la primera letra del apellido H. Lebesgue.

Así pues, $\|f\|_p$ es, de hecho, una seminorma (que no es una norma) en el espacio lineal $RL_p[a, b]$.

Una construcción análoga es válida también para los intervalos infinitos; los correspondientes espacios seminormalizados se designarán también mediante RL_p.

9. Analicemos un conjunto de todas las funciones continuas en el segmento $[a, b]$. Es un espacio lineal. Ya sabemos que se puede introducir en él la norma $\|f\|_C$ definida en el ejemplo 7 de este punto. Se puede también considerar en dicho espacio la seminorma (57.15), con la particularidad de que la seminorma (57.15) ya será en él una norma.

En efecto, si la función f es continua en el segmento $[a, b]$ y $\|f\|_p = 0$, $1 \leqslant p < +\infty$, y, por consiguiente,

$$\int_a^b |f(x)|^p dx = 0,$$

entonces de lo que la función $|f(x)|^p$, $x \in [a, b]$, es no negativa y continua se infiere (véase la propiedad 9 de la integral en el punto 28.1) que $f(x) \equiv 0$ en $[a, b]$.

Un espacio de funciones continuas en el segmento $[a, b]$ provisto de la norma (57.15) *se denota mediante* $CL_p[a, b]$.

De un modo semejante se construyen los espacios análogos para los intervalos no acotados, como también para las funciones de varias variables.

Si un mismo conjunto pertenece a diferentes espacios lineales seminormalizados o normalizados (por ejemplo, los espacios $C[a, b]$ y $CL_p[a, b]$ se componen de unas mismas funciones), entonces resulta útil con frecuencia estimar una norma (una seminorma) de estos elementos en términos de la otra. Los teoremas que expresan las estimaciones de esta índole se denominan *teoremas de encaje*.

Expliquemos lo dicho con un ejemplo enunciado en forma de un lema.

Lema 3. *Sea* $-\infty < a < b < +\infty$, $1 < p < +\infty$. *Si* $f \in RL_p[a, b]$, *se tiene*

$$\|f\|_1 \leqslant (b - a)^{1/q} \|f\|_p, \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{57.16}$$

y si $f \in RL_p[a, b] \cap S[a, b]$, *entonces*

$$\|f\|_p \leqslant (b - a)^{1/p} \|f\|_\infty. \tag{57.17}$$

DEMOSTRACIÓN. Teniendo presente que la seminorma $\|f\|_p$ se determina según la fórmula (57.15), obtenemos, haciendo uso de la desigualdad de Hölder (véase el p. 28.4*):

$$\|f\|_1 = \int_a^b |f(t)| \cdot 1 \, dt \leqslant$$

$$\leqslant \left[\int_a^b |f(t)|^p dt\right]^{1/p} \left[\int_a^b dx\right]^{1/q} = (b - a)^{1/q} \|f\|_p, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

con lo cual queda demostrada (57.16). La desigualdad (57.17) también se deduce directamente de las definiciones (57.14) y (57.15) de las normas correspondientes:

$$\|f\|_p = \left[\int_a^b |f(t)|^p dt\right]^{1/p} \leqslant \left\{\int_a^b \left[\sup_{[a, b]} |f(t)|\right]^p dt\right\}^{1/p} = \|f\|_\infty \left(\int_a^b dt\right)^{1/p}$$

$$= (b - a)^{1/p} \|f\|_\infty. \quad \square$$

Ejercicio 15. Denotemos con $C^1L_2[a, b]$ un subconjunto del espacio $CL_2[a, b]$, compuesto de las funciones continuamente derivables en el segmento $[a, b]$. Demuéstrese que

1) $C^1L_2[a, b]$ es un espacio lineal normalizado, si como norma de la función $f \in C^1L_2[a, b]$ se entiende su norma en el espacio $CL_2[a, b]$;

2) el operador de derivación D es un operador lineal no acotado $D: C^1L_2[a, b] \to CL_2[a, b]$.

Indicación: conviene examinar las funciones $\operatorname{sen} nx \in C^1L_2[-\pi, \pi]$.

57.5. PROPIEDADES DE LOS ESPACIOS SEMINORMALIZADOS

En los espacios seminormalizados se puede introducir el concepto de sucesión convergente y el de su límite.

Definición 25. *Si la sucesión de elementos $\{x_n\}$ de un espacio lineal seminormalizado (en particular, normalizado) X es de tal género que existe un elemento $x \in X$ tal que $\lim_{n\to\infty} \|x_n - x\| = 0$, entonces la sucesión $\{x_n\}$ se llama convergente en la seminorma (norma, respectivamente) hacia el elemento x y en este caso se escribe* $x = \lim_{n\to\infty} x_n$.

Introduciendo, en algún espacio lineal de funciones, diferentes seminormas (en particular, normas), obtendremos diferentes conceptos de convergencia de las sucesiones de funciones. Por ejemplo, la convergencia en el sentido de la norma (57.14) significa una convergencia uniforme; la convergencia en el sentido de la seminorma (57.15) ya es una convergencia de otro género; se denomina *convergencia en media* o, más detalladamente, en el sentido de p-media (a veces se habla simplemente de la convergencia en el sentido del espacio L_p). Ya nos encontramos con el caso particular de la convergencia de este género: para $p = 1$ (véanse el lema 2 en el p. 55.2, el corolario del lema 4 en el p. 56.7 y la métrica (57.2)), y para $p = 2$ (véase el corolario del teorema 12 del p. 55.9). Para $p = 2$ la convergencia en media se denomina también convergencia en el *sentido de la media cuadrática*.

Las desigualdades (57.16) y (57.17) entre diferentes seminormas de las funciones permiten establecer la relación entre diferentes tipos de convergencia de las sucesiones de funciones.

Por ejemplo, supongamos que la sucesión de funciones f_n, $n = 1, 2, \ldots$ y la función f son de tal género que

1°. La sucesión $\{f_n\}$ converge uniformemente en el segmento $[a, b]$ hacia la función f.

2°. Para todo $n = 1, 2, \ldots : f_n - f \in S[a, b] \cap RL_p[a, b]$.

Entonces la sucesión $\{f_n\}$ converge a la función f en el segmento $[a, b]$ también en el sentido de p-media, $1 \leqslant p < +\infty$.

En efecto, en virtud de (57.17), se verifica la desigualdad

$$\|f_n - f\|_p \leqslant (b - a)^{1/p} \|f_n - f\|_\infty.$$

La convergencia uniforme de la sucesión $\{f_n\}$ a la función f en el segmento $[a, b]$ significa que

$$\lim_{n\to\infty} \|f_n - f\|_\infty = 0.$$

Por consiguiente

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Ejercicio 16°. Constrúyase un ejemplo de sucesión de funciones continuas no negativas en un segmento que sea convergente en media, pero que no converja en ninguno de los puntos.

Cabe prestar atención en que en un espacio seminormalizado el límite para la sucesión convergente no es, en el caso general, único. Además, si $\lim_{n\to\infty} x_n = a$ y $\lim_{n\to\infty} x_n = b$, entonces la seminorma de la diferencia de dos límites es nula: $\|a - b\| = 0$. Esto se deduce inmediatamente de la desigualdad

$$\|a - b\| \leqslant \|a - x_n\| + \|x_n - b\|.$$

Lema 4. *Para cualesquiera dos elementos x e y de un espacio lineal seminormalizado X se verifica la desigualdad*

$$|\,\|x\| - \|y\|\,| \leqslant \|x - y\|. \tag{57.18}$$

DEMOSTRACIÓN. Puesto que

$$\|x\| = \|(x - y) + y\| \leqslant \|x - y\| + \|y\|,$$

se tiene

$$\|x\| - \|y\| \leqslant \|x - y\|$$

y, análogamente,

$$\|y\| - \|x\| \leqslant \|x - y\|.$$

De las dos últimas desigualdades proviene la desigualdad (57.18). □

Definición 26. *Sea X un espacio lineal seminormalizado (en particular, normalizado). Un conjunto $E \subset X$ se denomina acotado, o, más detalladamente, acotado en seminorma (en norma, respectivamente), si existe una constante $M > 0$ tal que para todo $x \in E$ se verifica la desigualdad $\|x\| \leqslant M$.*

Lema 5. *Si la sucesión $\{x_n\}$ converge en seminorma en X, es acotada.*

DEMOSTRACIÓN. Sea $x = \lim_{n\to\infty} x_n$; ya que la sucesión es convergente, existe tal n_0 que si $n \geqslant n_0$, entonces $\|x_n - x\| \leqslant 1$, y, por lo tanto

$$\|x_n\| = \|(x_n - x) + x\| \leqslant \|x_n - x\| + \|x\| \leqslant \|x\| + 1.$$

Pongamos $M = \max\{\|x_1\|, \|x_2\|, \ldots, \|x_{n_0-1}\|, \|x\| + 1\}$; en este caso, evidentemente, para todo $n = 1, 2, \ldots$ se verifica la desigualdad $\|x_n\| \leqslant M$. □

En un espacio lineal provisto de seminorma puede definirse el concepto de función continua. En lo que sigue (véase el p. 57.9) nos hará falta el concepto de continuidad de la función de una y dos variables en un espacio seminormalizado. Definamos estos conceptos.

Sea X un espacio seminormalizado. Una función f, real o compleja, definida en X, se llama *continua en el punto* $x_0 \in X$, si para cualquier $\varepsilon > 0$ existe tal $\delta > 0$ que para todos los $x \in X$, que satisfagan la condición $\|x - x_0\| < \delta$, se verifica la desigualdad

$$|f(x) - f(x_0)| < \varepsilon.$$

Sea Y también un espacio seminormalizado. Una función f, real o compleja, definida en el producto $X \times Y$, se denomina *continua en el punto* $(x_0, y_0) \in X \times Y$, si para cualquier $\varepsilon > 0$ existe tal $\delta > 0$ que para cualesquiera $(x, y) \in X \times Y$, que satisfagan las desigualdades $\|x - x_0\| < \delta$, $\|y - y_0\| < \delta$, se verifica la desigualdad

$$|f(x, y) - f(x_0 y_0)| < \varepsilon.$$

Si la función f es continua en todo punto de cierto conjunto, se llama *continua* en dicho conjunto.

Por supuesto, la definición de continuidad para los espacios seminormalizados puede enunciarse también con ayuda de sucesiones de los elementos del espacio.

Por ejemplo, una función numérica f, definida en el espacio seminormalizado X, se denomina continua en el punto x_0, si para toda sucesión $\{x_n\}$, que converge hacia x_0 en la seminorma del espacio X: $\lim_{n \to \infty} \|x_n - x_0\| = 0$ se verifica la igualdad $\lim_{n \to \infty} f(x_n) = f(x_0)$.

La equivalencia de las dos definiciones, enunciadas arriba, del límite de una función se demuestra según el mismo esquema que se ha usado en el caso en que X representaba un conjunto de números reales (véase el p. 5.7).

Lema 6. *La seminorma* $\|x\|$ *es una función continua en el espacio seminormalizado* X.

DEMOSTRACIÓN. Sean dados un elemento $x_0 \in X$ y un número $\varepsilon > 0$. En virtud del lema 4, para todos los x tales que $\|x - x_0\| < \varepsilon$, tenemos $|\|x\| - \|x_0\|| < \|x - x_0\| < \varepsilon$, es decir, la condición de continuidad de la función en X se cumple con la elección de $\delta = \varepsilon$. □

Definición 27. *Sean* X *e* Y *los espacios lineales seminormalizados (en particular, normalizados). Se denomina aplicación isomorfa o isomorfismo de los espacios seminormalizados (normalizados)* X *e* Y *una aplicación* f *que aplica de modo isomorfo el espacio* X, *en su calidad de espacio lineal, sobre el espacio* Y *(véase la definición* 19*) y que es de tal género que para cualquier* $x \in X$ *se verifica la igualdad*

$$\|x\|_X = \|f(x)\|_Y.$$

Si para los espacios lineales seminormalizados (normalizados) X *e* Y *existe una aplicación isomorfa de* X *sobre* Y, *dichos espacios se denominan isomorfos.*

Dos espacios isomorfos seminormalizados (normalizados) pueden diferir uno del otro sólo en la naturaleza de sus elementos y no en las propiedades del espacio. Por ello, en lo que sigue no vamos a distinguir a menudo los espacios isomorfos seminormalizados (normalizados) compuestos de diferentes elementos: tales espacios pueden "identificarse".

Expliquemos esto con detalles. Sean X e Y los espacios lineales seminormalizados, $Y \subset Y^*$, y sea $f: X \to Y$ una aplicación isomorfa. Consideraremos un conjunto

$X^* = X \cup (Y^* \setminus Y)$ que se obtiene del espacio X por adición a éste del conjunto $Y^* \setminus Y$. De este modo, $X^* \setminus X = Y^* \setminus Y$. Definamos para los elementos del conjunto X^* las operaciones de adición y multiplicación por un número y, además, la norma; se acompañarán ellas de índice X^*. Introduzcamos, con los fines de comodidad, la aplicación $F: X^* \to Y^*$, definida por la fórmula

$$F(x) \stackrel{\text{def}}{=} \begin{cases} f(x), & \text{si } x \in X, \\ x & \text{si } x \in X^* \setminus X. \end{cases} \tag{57.19}$$

Está claro que F es una aplicación biunívoca (biyección) del conjunto X^* sobre el Y^*.

Ahora, para cualesquiera $x \in X^*$, $y \in X^*$ y cualesquiera números λ, μ pongamos

$$(\lambda x + \mu y)_{X^*} \stackrel{\text{def}}{=} F^{-1}[\lambda F(x) + \mu F(y)],$$
$$\|x\|_{X^*} \stackrel{\text{def}}{=} \|F(x)\|.$$

El espacio X^*, definido de este modo, es lineal, seminormalizado (normalizado), isomorfo al espacio Y^* y contiene X en calidad de su subconjunto. Cuando decimos "identifiquemos en el espacio Y^* el conjunto Y con el espacio X isomorfo a Y", entendemos precisamente la consideración del espacio X^* mencionado anteriormente (compárese con la identificación de los espacios métricos isométricos en el p. 57.1).

Ejercicios. 17. Sea X un espacio lineal seminormalizado. Los elementos $x \in X$ e $y \in X$ se denominan *equivalentes*, si $\|x - y\| = 0$. Denotemos por $\hat{X}$ un conjunto cuyos elementos son las clases de elementos equivalentes del espacio X. Sean $\hat{x} \in \hat{X}$, $\hat{y} \in \hat{X}$, $x \in \hat{x}$, $y \in \hat{y}$, λ, un número cualquiera. Definamos $\hat{x} + \hat{y}$ como elemento de $\hat{X}$ que contiene $x + y$, y $\lambda\hat{x}$, como un elemento de $\hat{X}$ que contiene λx. Pongamos $\|\hat{x}\|_{\hat{X}} = \|x\|_X$. Demuéstrese que las definiciones dadas son correctas, es decir, no dependen de la elección de los elementos $x \in \hat{x}$ e $y \in \hat{y}$, y que $\hat{X}$ es un espacio lineal normalizado provisto de la norma $\|\hat{x}\|_{\hat{X}}$.

18. Demuéstrese que las funciones $x + y$ e λx son continuas en todo espacio lineal seminormalizado X (x e y son los elementos de este espacio y λ es un número), en otras palabras, que las operaciones de adición y multiplicación por un número son continuas en el espacio mencionado.

57.6. PROPIEDADES DE LOS ESPACIOS NORMALIZADOS

En el espacio lineal *normalizado* X puede introducirse, de modo natural, la distancia entre los elementos de dicho espacio. A saber, resulta válida la siguiente afirmación.

Lema 7. *El espacio lineal normalizado X es un espacio métrico provisto de la métrica*

$$\rho(x, y) = \|x - y\|, \tag{57.20}$$

con la particularidad de que la convergencia de las sucesiones en el espacio X en dicha métrica coincide con la convergencia en la norma.

DEMOSTRACIÓN. La función $\rho(x, y)$, definida según la fórmula (57.20), es realmente una distancia: las propiedades de la distancia (véase el p. 57.1) se deducen de las propiedades de la norma $1^\circ - 4^\circ$ (compruébese). La segunda afirmación del lema es obvia. □

Diremos que la *métrica* (57.20) *se engendra por la norma dada del espacio X*. Por ejemplo, una métrica engendrada por la norma $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$ en un espacio lineal aritmético de los vectores reales n-dimensionales $x = (x_1, x_2, \ldots, x_n)$ es la métrica del espacio euclídeo R^n definida mediante la fórmula (18.1).

Una sucesión de puntos del espacio X, fundamental respecto de la métrica (57.20), se llama *fundamental respecto de la norma* dada en el espacio X.

Ejercicio 19. Demuéstrese que un conjunto en el espacio lineal normalizado está acotado en norma (véase la definición 26 en el p. 57.5) cuando, y sólo cuando, está acotado como conjunto de un espacio métrico en el sentido de la métrica (57.20) (véase el ejercicio 1 en el p. 57.1).

Ejemplo. Examinemos el espacio l_p de sucesiones de números reales con la norma (57.10). Designemos mediante $\{e_n\}$ una sucesión cuyo n-ésimo término es igual a la unidad, mientras que todos los términos restantes son nulos. Es evidente que para $n \neq m$

$$\|e_n - e_m\| = (1 + 1)^{1/p} = 2^{1/p}.$$

Por eso la sucesión de elementos e_n, $n = 1, 2, \ldots$, del espacio l_p no puede contener una subsucesión fundamental ni tampoco, por consiguiente, una que sea convergente.

La sucesión $\{e_n\}$ está acotada, pues para todo n tenemos $\|e_n\| = 1$. Forma un conjunto cerrado en l_p, puesto que el conjunto $\{e_n\}$ no tiene puntos límites en l_p (de lo contrario, en la sucesión se encontraría una subsucesión convergente).

De este modo, en un espacio de dimensión finita existen sucesiones acotadas, de las cuales no se puede separar una que sea convergente. Existen también conjuntos cerrados acotados en los cuales no toda sucesión de sus puntos permite que de ella sea separada una convergente.

OBSERVACIÓN 1. Si en el espacio lineal X están definidas dos normas de elementos $\|\cdot\|^{(1)}$ y $\|\cdot\|^{(2)}$ y estas normas son equivalentes (véase la definición 23 en el p. 57.4), entonces la sucesión $x_n \in X$, $n = 1, 2, \ldots$, converge hacia el elemento $x \in X$ en el sentido de la norma $\|\cdot\|^{(1)}$ cuando, y sólo cuando, converge a x en el sentido de la norma $\|\cdot\|^{(2)}$.

En efecto, en vista de que las normas $\|\cdot\|^{(1)}$ y $\|\cdot\|^{(2)}$ son equivalentes, existen unas constantes $c_1 > 0$ y $c_2 > 0$ tales que se verifican las desigualdades

$$c_1\|x_n - x\|^{(2)} \leqslant \|x_n - x\|^{(1)} \leqslant c_2\|x_n - x\|^{(2)}.$$

De estas desigualdades deriva directamente la equivalencia de las convergencias de la sucesión $\{x_n\}$ hacia x en el sentido de las normas $\|\cdot\|^{(1)}$ y $\|\cdot\|^{(2)}$.

De la equivalencia, demostrada en el teorema 2 del p. 57.4, de todas las normas en un espacio de dimensión finita se infiere que las convergencias de las sucesiones de sus puntos en todas las normas son equivalentes. Por cuanto la convergencia en norma cuadrática $\|x\|_2$ es equivalente a la convergencia en coordenadas (véanse los pp. 18.1 y 18.4), la convergencia de la sucesión de puntos en un espacio de dimensión finita en cualquier norma es equivalente a la convergencia de las sucesiones numéricas de las coordenadas de los puntos en consideración respecto de una base arbitraria.

OBSERVACIÓN 2. Observemos que en el caso cuando una seminorma no es norma, incluso una función simple como la lineal, en un espacio seminormalizado lineal de dimensión finita puede resultar no continua. Consideremos, por ejemplo, un espacio aritmético bilineal X de los vectores $x = (x_1, x_2)$ con la seminorma $\|x\| = |x_1|$. Es realmente una seminorma, puesto que $\|x\| = |x_1| \geqslant 0$. Además, para cualquier número λ se tiene $\lambda x = (\lambda x_1, \lambda x_2)$, por lo cual $\|\lambda x\| = |\lambda x_1| = |\lambda|\,|x_1| = |\lambda|\,\|x\|$. Por fin, si $y = (y_1, y_2)$ es también un elemento de X, entonces $x + y = (x_1 + y_1, x_2 + y_2)$, por consiguiente, $\|x + y\| = |x_1 + y_1| \leqslant |x_1| + |y_1| = \|x\| + \|y\|$. De este modo, todas las propiedades de la seminorma están cumplidas.

Mostremos que la función lineal $f(x) = x_2$ no es continua en X. En efecto, para la sucesión $x^{(n)} = (1/n, 1)$ cualquier punto del tipo $x = (0, x_2)$ (x_2 es arbitrario) es el límite para dicha sucesión en el sentido de la seminorma en consideración:

$$\lim_{n \to \infty} \|x^{(n)} - x\| = \lim_{n \to \infty} \frac{1}{n} = 0.$$

En particular, el punto $O = (0, 0)$ es el límite de la sucesión $\{x^{(n)}\}$. Sin embargo,

$$\lim_{n \to \infty} f(x^{(n)}) = 1 \neq 0 = f(O).$$

Esto es indicio de que la función $f(x) = x_2$ no es continua en la seminorma $\|x\| = |x_1|$.

Subrayemos, no obstante, que si en un espacio de dimensión finita la seminorma es una norma, entonces toda función lineal será continua respecto de esta norma. Efectivamente, sea X un espacio lineal normalizado n-dimensional y sea f una funcional lineal en X. Supongamos que $\{e_1, \ldots, e_n\}$ es una base en X y, por consiguiente, todo elemento $x \in X$ es representable, y además, de manera unívoca, en la forma $x = x_1 e_1 + \ldots + x_n e_n$. Por cuanto f es una funcional lineal, se tiene

$$f(x) = f(x_1 e_1 + \ldots + x_n e_n) = x_1 f(e_1) + \ldots + x_n f(e_n) = a_1 x_1 + \ldots + a_n x_n,$$

donde $a_k = f(e_k)$, $k = 1, 2, \ldots, n$ son los números fijos para f. Recordando que la convergencia de una sucesión de puntos en cualquier norma en un espacio de dimensión finita es equivalente a la convergencia de dicha sucesión en coordenadas, nos convencemos en seguida de que la continuidad de la función f deriva realmente de la fórmula obtenida $f(x) = a_1 x_1 + \ldots + a_n x_n$.

Lema 8. *La norma es una función continua en un espacio lineal normalizado en el sentido de la métrica* (57.20).

En virtud de la igualdad (57.20), esta afirmación se deduce de lo que una seminorma es continua respecto de otra seminorma (véase el lema 6 en el p. 57.5).

Definición 28. *Un espacio lineal normalizado se denomina completo, si es espacio métrico completo en el sentido de la métrica engendrada por la norma de dicho espacio.*

*Todo espacio lineal normalizado completo se llama espacio de Banach**).

---

*) S. Banach (1892 — 1945), matemático polaco.

El espacio lineal normalizado $C[a, b]$ de funciones continuas en el segmento $[a, b]$ dotado de la norma (57.14) es un espacio de Banach. Nos hemos convencido de esto en el p. 57.1, al considerar un espacio métrico de funciones continuas en el segmento $[a, b]$ con la distancia (57.1) la que se engendra precisamente por la norma (57.14). Hemos visto que la completitud del espacio $C[a, b]$ se desprende de lo que la convergencia de la sucesión en este espacio significa la *convergencia uniforme* de la misma en el segmento $[a, b]$.

Teorema 3. *Todo espacio lineal normalizado está contenido, y además, es denso en cierto espacio de Banach.*

DEMOSTRACIÓN. De acuerdo con el teorema 1 en el p. 57.1, es suficiente mostrar que a la completación X^* de un espacio lineal normalizado X (considerado como un espacio métrico provisto de la métrica (57.20)) pueden ser prolongadas desde X las operaciones algebraicas y la norma. Es posible hacerlo con ayuda del paso al límite. Al igual que en la demostración del teorema 1, convengamos en considerar que $X \subset X^*$, en otras palabras, identifiquemos el espacio X con un subespacio, isométrico a éste, de la completación X^* construida en X.

Sea, por ejemplo, $x \in X^*$ e $y \in X^*$. Siendo X denso, existen en X^* las sucesiones $x_n \in X$ e $y_n \in X$, $n = 1, 2, \ldots$, tales que $\lim\limits_{n \to \infty} x_n = x$, $\lim\limits_{n \to \infty} y_n = y$.

Probemos que la sucesión $\{x_n + y_n\}$ converge. En efecto,

$$\rho(x_n + y_n, x_m + y_m) = \|(x_n + y_n) - (x_m + y_m)\| \leqslant$$
$$\leqslant \|x_n - x_m\| + \|y_n - y_m\| = \rho(x_n, x_m) + \rho(y_n, y_m).$$

De la convergencia de las sucesiones $\{x_n\}$ e $\{y_n\}$ se desprende que ellas son fundamentales, por lo cual la sucesión $\{x_n + y_n\}$ es también fundamental y, por lo tanto, siendo X^* completo, convergente.

Pongamos, por definición,

$$x + y = \lim_{n \to \infty} (x_n + y_n).$$

Análogamente se determina también, pasando al límite, λx, $x \in X^*$.

Es *fácil comprobar que las operaciones algebraicas* $x + y$, λx, definidas de esta forma para los elementos de la completación X^*, no dependen de cómo se eligen las sucesiones $\{x_n\}$ e $\{y_n\}$ tales que $x_n \to x$, $y_n \to y$, $x_n \in X$, $y_n \in X$, $n = 1, 2, \ldots$. Tampoco es difícil convencerse de que en el caso cuando los elementos pertenezcan al espacio de partida X, las operaciones algebraicas, definidas por nosotros, coinciden con las dadas.

Determinemos ahora la norma para $x \in X^*$. Sea $x_n \in X$, $n = 1, 2, \ldots$, y $\lim\limits_{n \to \infty} x_n = x$. Mostremos que la sucesión numérica $\{\|x_n\|\}$ es fundamental. En efecto, de la desigualdad (57.18) tenemos, para cualesquiera n y m naturales:

$$|\|x_n\| - \|x_m\|| \leqslant \|x_n - x_m\| = \rho(x_n, x_m). \tag{57.21}$$

La sucesión $\{x_n\}$, siendo convergente, es, además, fundamental, por lo cual de la desigualdad (57.21) se deduce que la sucesión numérica $\{\|x_n\|\}$ es también fundamental y, por lo tanto, convergente.

Pongamos, por definición,

$$\|x\| = \lim_{n \to \infty} \|x_n\|.$$

La norma $\|x\|$, $x \in X^*$, definida de esta manera, no depende de cómo se elige la sucesión $x_n \in X$, $n = 1, 2, \ldots$, tal que $x_n \to x$. Es fácil comprobar también, pasando al límite, que para la función $\|x\|$, $x \in X^*$, se cumplen las propiedades de la norma $1^\circ - 4^\circ$ y que para $x \in X$ obtenemos la norma anterior. □

A título de ejemplo indiquemos el espacio lineal normalizado $CL_p[a, b]$ de funciones continuas en el segmento $[a, b]$ provisto de la norma (57.15). Cuando $p = 1$, esta norma engendra la métrica (57.2). Se puede mostrar que un espacio métrico de funciones continuas provisto de la métrica (57.2) no es completo. De acuerdo con el teorema demostrado, dicho espacio lineal normalizado de funciones continuas en el segmento $[a, b]$ puede ser completado hasta que se obtenga un espacio completo. Este espacio de Banach se designa mediante $L[a, b]$.

Definición 29. *Un sistema de elementos x_α, $\alpha \in \mathfrak{A}$ ($\mathfrak{A}$ es cierto conjunto de índices), del espacio lineal seminormalizado X se denomina completo en dicho espacio, si para todo elemento $x \in X$ y cada número $\varepsilon > 0$ existen tales elementos $x_{\alpha_1}, \ldots, x_{\alpha_n}$ del sistema dado y tales números $\lambda_1, \ldots, \lambda_n$ que se verifica la desigualdad*

$$\|x - (\lambda_1 x_{\alpha_1} + \ldots + \lambda_n x_{\alpha_n})\| < \varepsilon. \tag{57.22}$$

Enunciemos esta definición en una forma algo diferente, introduciendo previamente una noción más.

Definición 30. *Un conjunto $A \subset X$ se llama denso en el espacio seminormalizado X, si para todo elemento $x \in X$ y todo $\varepsilon > 0$ se encontrará tal elemento $a \in A$ se verifique*

$$\|x - a\| < \varepsilon.$$

Si X es un espacio normalizado y, por lo tanto, métrico, entonces, en virtud de (57.20), la definición 30 conduce a la misma noción de densidad de un conjunto que la definición 6 del p. 57.1. Ahora podemos decir:

Un sistema $\{x_\alpha\}$, $\alpha \in \mathfrak{A}$, es completo en el espacio X, si el conjunto de todas las combinaciones lineales de sus elementos, es decir, su cápsula lineal (véase la definición 15 en el p. 57.2) *forma un conjunto denso en X.*

Si X es un espacio normalizado, tiene sentido en él, como en cualquier otro espacio métrico, el concepto de clausura de un conjunto, y por cuanto la densidad de cierto conjunto en un espacio métrico significa que la clausura del conjunto citado coincide con el mismo espacio (véase la definición 6 en el p. 57.1), entonces la definición 30 en este caso puede parafrasearse del modo siguiente:

un sistema de elementos x_α, $\alpha \in \mathfrak{A}$ ($\mathfrak{A}$ es cierto conjunto de índices), del espacio lineal normalizado X se llama completo, si la clausura de su cápsula lineal (véase el p. 57.2) *coincide con todo el espacio X.*

Con el caso particular del concepto de completitud para un sistema de funciones ya nos hemos encontrado en el p. 55.8.

Definición 31. *Si en el espacio lineal normalizado X existe un conjunto numerable de elementos que forma un sistema completo del espacio X, este último se denomina separable.*

En conclusión de este punto introduzcamos el concepto de base, y, ante todo, el de serie en el espacio X.

Definición 32. *Sea x_n, $n = 1, 2, \ldots$, una sucesión de elementos del espacio lineal normalizado X. Pongamos $s_n = x_1 + \ldots + x_n$, $n = 1, 2, \ldots$; un par de sucesiones $\{x_n\}$, $\{s_n\}$ se llama serie (con el término general x_n) y se designa*

$$\sum_{n=1}^{\infty} x_n; \tag{57.23}$$

los elementos s_n se denominan n-ésimas sumas parciales de la serie (57.23).

Si la sucesión s_n, $n = 1, 2, \ldots$, converge en el espacio X, la serie (57.23) se llama *convergente*. En este caso el límite $s = \lim_{n \to \infty} s_n$ de la sucesión s_n, $n = 1, 2, \ldots$, se llama *suma de la serie* (57.23) y se escribe

$$\sum_{n=1}^{\infty} x_n = s.$$

De este modo, al igual que en el caso de las series numéricas, el mismo símbolo $\sum_{n=1}^{\infty} x_n$ se empleará para designar tanto la propia serie, como también su suma, si la serie converge.

Lo mismo que para las series numéricas, para las series en los espacios lineales normalizados subsisten las siguientes afirmaciones.

Si la serie (57.23) *converge, converge también la serie $\sum_{n=1}^{\infty} \lambda x_n$, con la particularidad de que si $\sum_{n=1}^{\infty} x_n = s$, entonces $\sum_{n=1}^{\infty} \lambda x_n = \lambda s$.*

Si en el espacio X convergen dos series, converge también una serie cuyo término es igual a la suma de sus términos generales y la suma de esta última serie es igual a la suma de las sumas de las series dadas.

Definición 33. *La sucesión de elementos e_n, $n = 1, 2, \ldots$ de un espacio lineal normalizado se llama base, si, cualquiera que sea el elemento x, existe una sucesión, y sólo una, de los números λ_n, $n = 1, 2, \ldots$, tal que*

$$x = \sum_{n=1}^{\infty} \lambda_n e_n. \tag{57.24}$$

De este modo, si la sucesión $\{e_n\}$ es una base del espacio X, para todo elemento $x \in X$ existe una, y sólo una, sucesión de los números $\{\lambda_n\}$ tal que para cualquier $\varepsilon > 0$ se tiene tal número n_ε que para todos los $n \geqslant n_\varepsilon$ se verifica la desigualdad

$$\|x - (\lambda_1 e_1 + \ldots + \lambda_n e_n)\| < \varepsilon. \tag{57.25}$$

La fórmula (57.24) se denomina desarrollo del elemento x según la base $\{e_n\}$.

No es difícil convencerse de que si el sistema de elementos $\{e_n\}$ forma una base, es linealmente independiente. Esto proviene en seguida de la unicidad del desarrollo de los elementos del espacio según la base. En efecto, si los elementos e_n, $n = 1, 2,$

. . . , resultaran ser linealmente dependientes, se encontraría entre ellos un conjunto finito de tales $e_{n_1}, \ldots, e_{n_k}$, que para ciertos números $\lambda_1, \ldots, \lambda_k$, no todos iguales a cero, tendría lugar la igualdad $\lambda_1 e_{n_1} + \ldots + \lambda_k e_{n_k} = 0$, es decir, se obtendría un desarrollo del cero según los elementos de la base con coeficientes que no son todos nulos. Por cuanto para el cero se tiene un desarrollo trivial $0 = \sum_{n=1}^{\infty} 0 e_n$, queda con esto perturbada la condición de unicidad del desarrollo de los elementos según la base.

Si un espacio lineal normalizado dispone de una base compuesta por un conjunto finito o numerable de elementos, este espacio es separable. En efecto, no es difícil comprobar que el conjunto de todas las combinaciones lineales finitas de elementos de las bases citadas con coeficientes racionales es numerable y denso en todo el espacio.

OBSERVACIÓN. Recalquemos la diferencia que existe entre la sucesión de elementos que forman un sistema completo y la sucesión de elementos que forman una base. En el primer caso los coeficientes λ_k, $k = 1, 2, \ldots, n$, en la desigualdad (57.22) dependen, en el caso general, no sólo de la elección del elemento $x \in X$, sino también de la elección del número ε. Por el contrario, en el segundo caso los coeficientes λ_k, $k = 1, 2, \ldots, n$, en la desigualdad (57.25) se determinan sólo por el mismo elemento (se llaman *coeficientes del desarrollo del elemento x según la base dada*, o bien *coordenadas del elemento x para la base dada*) y es sólo su cantidad, es decir, el número n_ε, el que depende de la elección de ε.

Existen espacios de Banach separables en los que no hay base. En el punto que sigue será considerada una clase más estrecha de los espacios, donde la base siempre existe.

57.7. ESPACIOS LINEALES PROVISTOS DE PRODUCTO ESCALAR

Definición 34. *Una función real, definida en el conjunto de los pares ordenados de elementos de un espacio lineal real y denotada por* (x, y), $x \in X$, $y \in X$, *se llama multiplicación escalar, si satisface las siguientes condiciones*:

1°) $(x, y) = (y, x)$, $x \in X$, $y \in X$;

2°) $(\lambda x + \mu y, z) = \lambda(x, z) + \mu(y, z)$, $x \in X$, $y \in X$, $z \in X$, λ y μ *son unos números reales*;

3°) $(x, x) \geqslant 0$, $x \in X$;

4°) *si* $(x, x) = 0$, *entonces* $x = 0$.

Observemos que de la propiedad 2° se desprende que para todo $x \in X$ se verifica la igualdad

$$(x, 0) = 0.$$

En efecto, $(x, 0) = (x, 0 \cdot 0) = 0(x, 0) = 0$.

Definición 35. *Una función real* (x, y), *definida en el conjunto de los pares ordenados de un espacio lineal real* X, $x \in X$, $y \in X$, *y que satisface sólo las condiciones* 1°, 2°, 3°, *se llama multiplicación semiescalar.*

De un modo análogo se introduce también el concepto de multiplicación semiescalar (en particular, escalar) en el espacio lineal complejo R. En este caso la función

de valores complejos (x, y) lleva el nombre de multiplicación semiescalar (escalar respectivamente), si satisface la propiedad 2° para cualesquiera números complejos λ y μ, la propiedad 3° y la propiedad

$$1'^{\circ}) (x, y) = \overline{(y, x)}, \quad x \in X, \quad y \in X,$$

donde, como siempre, la raya por encima del número significa un número complejo conjugado de él.

En adelante, por espacio lineal se entenderá un espacio lineal real, si no se especifica alguna otra cosa.

El resultado de una multiplicación escalar (semiescalar) de dos elementos $x \in X$ e $y \in Y$ se denomina *producto escalar (semiescalar)* (x, y). Los espacios lineales, para cuyos elementos está definida la operación de multiplicación escalar (semiescalar), se llaman *espacios lineales provistos de producto escalar (semiescalar)*.

Lema 9. *Para cualquier par de vectores x e y de un espacio lineal X provisto de producto semiescalar se verifica la desigualdad*

$$(x, y)^2 \leqslant (x, x)(y, y), \tag{57.26}$$

llamada desigualdad de Cauchy — Schwartz.

DEMOSTRACIÓN. Para todo número real λ, en virtud de la propiedad 3° de la multiplicación semiescalar, se tiene

$$(\lambda x + y, \lambda x + y) \geqslant 0.$$

Al aplicar las propiedades 1° y 2° de la multiplicación semiescalar, obtendremos:

$$\lambda^2 (x, x) + 2\lambda (x, y) + (y, y) \geqslant 0.$$

Si $(x, x) = 0$, entonces $2\lambda (x, y) + (y, y) \geqslant 0$. Por cuanto esto es cierto para cualquier λ real, $(x, y) = 0$, y, por consiguiente, la desigualdad (57.26) es lícita: sus ambos miembros se reducen a cero. En cambio, si $(x, x) \neq 0$, entonces el discriminante del trinomio obtenido, cuadrático respecto de λ, es no positivo:

$$(x, y)^2 - (x, x)(y, y) \leqslant 0,$$

lo que es equivalente a la condición (55.26).

Corolario. *Para cualquier par de vectores de un espacio lineal provisto de producto semiescalar se verifica la desigualdad*

$$\sqrt{(x + y, x + y)} \leqslant \sqrt{(x, x)} + \sqrt{(y, y)}, \quad x \in X, \quad y \in X.$$

En efecto, al aplicar la desigualdad de Cauchy — Schwartz, obtendremos:

$$(x + y, x + y) = (x, x) + 2(x, y) + (y, y) \leqslant$$

$$\leqslant (x, x) + 2\sqrt{(x, x)(y, y)} + (y, y) = [\sqrt{(x, x)} + \sqrt{(y, y)}]^2. \quad \square$$

Ejercicio 20. Demuéstrese que en el espacio lineal complejo X provisto de producto semiescalar se verifica la desigualdad

$$|(x, y)|^2 \leqslant (x, x)(y, y), \quad x \in X, \quad y \in X$$

Si en un espacio lineal X provisto de producto semiescalar se pone

$$\|x\| = \sqrt{(x, x)}, \quad x \in X, \tag{57.27}$$

entonces la función $|x|$ satisface las propiedades 1° — 3° de la seminorma. La propiedad 1° de la seminorma proviene de la propiedad 3°, de la multiplicación semiescalar, la propiedad 2°, de la propiedad 2° y la propiedad 3° de la seminorma, del corolario del lema 9.

Cuando la multiplicación semiescalar es escalar, la seminorma (57.27) es una norma. En efecto, la propiedad 4° de la norma deriva de la propiedad 4° de la multiplicación escalar. De este modo, se ha demostrado la siguiente afirmación.

Lema 10. *Todo espacio lineal provisto de producto escalar (semiescalar, respectivamente) es un espacio normalizado (seminormalizado, respectivamente) con una norma (seminorma, respectivamente) determinada por la fórmula* (57.27) *y, por consiguiente, un espacio métrico con la métrica* (57.20).

La seminorma (57.27) se llamará *seminorma (norma,* respectivamente) *engendrada por el producto semiescalar (escalar) dado.* La distancia (57.20), engendrada por la norma (57.27) de un espacio lineal dotado de producto escalar, se llamará también *distancia engendrada por el producto escalar dado.*

Al aplicar la designación de la seminorma, podemos escribir la desigualdad (57.26) en la forma

$$|(x, y)| \leqslant \|x\| \, \|y\|. \tag{57.28}$$

57.8. EJEMPLOS DE ESPACIOS LINEALES PROVISTOS DE PRODUCTO ESCALAR

1. En el conjunto de números reales R una operación ordinaria de multiplicación es también multiplicación escalar en el sentido de la definición 34.

En el conjunto de números complejos C *el producto escalar* de los números x e y está representado por el producto $x\bar{y}$.

2. El espacio vectorial real aritmético n-dimensional R^n, donde el producto escalar de los vectores $x = (x_1, \ldots, x_n) \in R^n$ e $y = (y_1, \ldots, y_n) \in R^n$ se determina según la fórmula (véase (18.32) en el p. 18.4)

$$(x, y) = x_1 y_1 + \ldots + x_n y_n,$$

es un espacio lineal provisto de producto escalar en el sentido de la definición 34 del p. 57.7. En este caso la norma del elemento $x \in R^n$ coincide con su longitud $|x|$ (véase el p. 57.4, ejemplo 2):

$$\|x\| = |x| = \sqrt{x_1^2 + \ldots + x_n^2},$$

y la métrica correspondiente, con la distancia en el espacio puntual aritmético n-dimensional:

$$\rho(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

Recordemos que para este espacio la desigualdad de Cauchy — Schwartz ha sido demostrada antes (véanse el lema 1 en el p. 18.1 y la desigualdad (18.39) en el p. 18.4).

En el espacio complejo aritmético C^n (véase el p. 57.2) el producto escalar se introduce por medio de la fórmula

$$(x, y) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n, x = (x_1, \ldots, x_n) \in C^n, \quad y = (y_1, \ldots, y_n) \in C^n.$$

3. Analicemos el espacio lineal seminormalizado $RL_2[a, b]$ (del ejemplo 8, el p. 57.4), compuesto de las funciones de cuadrado integrable (por regla general, en el sentido impropio) en el segmento $[a, b]$, es decir, de tales funciones f, para las cuales

$$\int_a^b f^2(t)dt < +\infty.$$

Sea $f \in RL_2[a, b]$ y $g \in RL_2[a, b]$. Recordemos que el producto de las funciones que son integrables según Riemann en cierto segmento es también integrable según Riemann en dicho segmento. Por eso, en cualquier segmento $[\xi, \eta] \subset [a, b]$ que no contiene puntos singulares de las funciones f y g (véase el p. 55.1), el producto fg es también integrable según Riemann y, por lo tanto, hay sentido en considerar la integral impropia

$$\int_a^b f(t)g(t)dt. \tag{57.29}$$

Por cuanto, además, en cualquier punto x, que no sea singular para las funciones f y g, es válida la desigualdad

$$|f(t)g(t)| \leqslant \frac{f^2(t) + g^2(t)^{*)}}{2},$$

entonces la integral (57.29) converge y, además, absolutamente.

El producto semiescalar en este espacio se determina por la fórmula

$$(f, g) = \int_a^b f(t)g(t)dt. \tag{57.30}$$

Las propiedades 1°, 2°, 3° del producto semiescalar se comprueban con facilidad. El espacio obtenido, provisto de producto semiescalar (57.30), se designará también mediante $RL_2[a, b]$.

Observemos que la desigualdad (57.26) en este caso puede ser escrita en la forma siguiente:

$$\left(\int_a^b f(t)g(t)dt\right)^2 \leqslant \int_a^b f^2(t)dt \int_a^b g^2(t)dt;$$

representando en sí un caso particular de la desigualdad de Hölder (véase el p. 28.4*) para $p = q = 2$ y llevando el nombre de desigualdad de Cauchy — Buniakovski**).

---

*) Dicha desigualdad proviene de la siguiente desigualdad evidente: $(|f(t)| - |g(t)|)^2 \geqslant 0$.

**) V. Ya. Buniakovski (1804 — 1889), matemático ruso.

La seminorma engendrada por el producto semiescalar (57.30) tiene, evidentemente, la forma

$$\|f\| = \sqrt{\int_a^b f^2(t)dt}, \tag{57.31}$$

es decir, coincide con la seminorma (57.15) examinada en el ejemplo 8, del p. 57.4, para $p = 2$. De aquí se deduce que el producto semiescalar (57.30) no es escalar, puesto que en el p. 57.4 se ha establecido que la seminorma (57.15) no es una norma, para todo $p \geqslant 1$.

Sin embargo, en el subespacio $CL_2[a, b]$ del espacio $RL_2[a, b]$, que se compone sólo de las funciones continuas en el segmento $[a, b]$, el producto semiescalar (57.30) ya es escalar, pues, según se ha demostrado en el ejemplo 9 del p. 57.4, en este caso

$$\|f\|_2 = \sqrt{(f, f)}, \quad f \in CL_2[a, b]$$

es no sólo una seminorma, sino también una norma.

Para la distancia entre dos funciones continuas f y g en este espacio obtenemos la fórmula

$$\rho(f, g) = \|f - g\| = \left\{ \int_a^b [f(x) - g(x)]^2 dx \right\}^{1/2} \tag{57.32}$$

Ya nos hemos encontrado con la convergencia de las funciones en el sentido de esta métrica: véase, por ejemplo, el corolario del teorema 12 en el p. 55.9.

Todo lo dicho se extiende de modo natural a las funciones definidas en cualquier intervalo infinito, en particular, en todo el eje.

Ejercicio 21. Sea X un espacio lineal provisto de producto semiescalar. Los elementos $x \in X$ e $y \in Y$ se llaman *equivalentes*, si $\|x - y\|^2 = (x - y, x - y) = 0$. Designemos mediante $\hat{X}$ un conjunto cuyos elementos están constituidos por las clases de elementos equivalentes del espacio X. Sea $\hat{x} \in \hat{X}$, $\hat{y} \in \hat{X}$, $x \in \hat{x}$, $y \in \hat{y}$, sean λ y μ unos números. Definamos $\lambda\hat{x} + \mu\hat{y}$ como un elemento del conjunto $\hat{X}$ que contiene $\lambda x + \mu y$, y pongamos $(\hat{x}, \hat{y}) = (x, y)$. Demuéstrese que estas definiciones son correctas, es decir, no dependen de la elección de los elementos $x \in \hat{x}$ e $y \in \hat{y}$, y que $\hat{X}$ *es un espacio lineal, mientras que* $(\hat{x}, \hat{y})$ es el producto escalar en $\hat{X}$.

57.9. PROPIEDADES DE LOS ESPACIOS LINEALES PROVISTOS DE PRODUCTO ESCALAR. ESPACIOS DE HILBERT

Un espacio lineal provisto de producto semiescalar es también, de acuerdo con (57.27), seminormalizado. Por eso, para él están definidos el concepto de sucesión convergente, de su límite y el de función continua (véase el p. 57.5).

Lema 11. *El producto semiescalar* (x, y) *es una función continua* (véase el p. 57.5) *de sus argumentos* x e y *en el conjunto* $X \times X$, $x \in X$, $y \in X$.

DEMOSTRACIÓN. Efectivamente, para cualesquiera $x_0 \in X$, $y_0 \in X$, $x \in X$ e $y \in X$ se verifica la desigualdad

$$|(x_0, y_0) - (x, y)| = |(x_0 - x, y_0)| + (x, y_0 - y)| \leqslant$$
$$\leqslant \|x_0 - x\| \|y_0\| + \|x\| \|y - y_0\|, \tag{57.33}$$

de la cual proviene inmediatamente la citada continuidad del producto semiescalar. En efecto, si $x \in U(x_0, \delta)$, $y \in U(y_0, \delta)$, entonces, al observar que $\|x\| \leqslant \|x - x_0\| + \|x_0\| < \|x_0\| + \delta$, de (57.33) obtenemos

$$|(x_0, y_0) - (x, y)| < \delta\|y_0\| + (|x_0| + \delta)\delta.$$

De aquí se infiere que para cualquier $\varepsilon > 0$ fijo siempre podemos elegir $\delta = \delta(\varepsilon) > 0$ de modo tal que para $x \in U(x_0, \delta)$, $y \in U(y_0, \delta)$ se verifique la desigualdad $|(x_0, y_0) - (x, y)| < \varepsilon$; para ello basta elegir $\delta > 0$ de tal manera que sea $\delta\|y_0\| + (\|x_0\| + \delta)\,\delta < \varepsilon$, lo que, evidentemente, es siempre posible. □

En el espacio X provisto de producto semiescalar podemos hablar sobre la convergencia de las series en seminorma engendrada por el producto semiescalar: la serie $\sum_{n=1}^{\infty} x_n, x_n \in X, x = 1, 2, \ldots$, se llama *convergente*, si la sucesión de sus sumas parciales $s_n = \sum_{k=1}^{n} x_k$ converge en la seminorma mencionada hacia cierto elemento $s \in X$, el cual se denomina suma de la serie: $s = \sum_{n=1}^{\infty} x_n$.

Indiquemos que la suma de una serie en el espacio provisto de producto semiescalar no está definida unívocamente. No obstante, si $s = \sum_{n=1}^{\infty} x_n$ y $s^* = \sum_{n=1}^{\infty} x_n$, es decir, si s y s^* son las sumas de una misma serie, entonces $\|s^* - s\| = 0$ (véase el p. 57.5), por lo cual para cualquier elemento $a \in X$ tiene lugar la igualdad $(s^*, a) = (s, a)$. En efecto, en virtud de la desigualdad de Cauchy — Schwartz, para un producto semiescalar tenemos

$$|(s^*, a) - (s, a)| = |(s^* - s, a)| \leqslant \|s^* - s\|\,\|a\| = 0.$$

De la continuidad del producto semiescalar en todo el espacio se desprende, por ejemplo, que las series en un espacio provisto de producto escalar pueden multiplicarse término a término no sólo por los factores numéricos, sino también por los elementos del mismo espacio. Demostrémoslo.

Lema 12. *Supongamos que en el espacio X provisto de producto semiescalar está dada una serie convergente*

$$\sum_{n=1}^{\infty} x_n = s, \quad x_n \in X, \quad n = 1, 2, \ldots .$$

Entonces para cada elemento $a \in X$ una serie numérica que se obtiene de la dada multiplicándola término a término por a es también convergente y

$$\sum_{n=1}^{\infty} (x_n, a) = (s, a).$$

En otras palabras, para la serie convergente $\sum_{n=1}^{\infty} x_n$ y cualquier elemento $a \in X$ se verifica la desigualdad

$$\sum_{n=1}^{\infty}(x_n, a) = \left(\sum_{n=1}^{\infty} x_n, a\right).$$

DEMOSTRACIÓN. Por cuanto

$$s = \lim_{n\to\infty}\sum_{k=1}^{n} x_k,$$

se tiene

$$\sum_{n=1}^{\infty}(x_n, a) = \lim_{n\to\infty}\sum_{k=1}^{n}(x_k, a) = \lim_{n\to\infty}\left(\sum_{k=1}^{n} x_k, a\right) = \left(\lim_{n\to\infty}\sum_{k=1}^{n} x_k, a\right) = (s, a).$$

Ejemplo. Consideraremos el espacio $RL_2[a, b]$ del ejemplo 3 en el p. 57.8. Supongamos que la serie $\sum_{n=1}^{\infty} f_n(t)$ de funciones $f_n \in RL_2[a, b]$ converge en este espacio a la función $f \in RL_2[a, b]$:

$$\sum_{n=1}^{\infty} f_n(t) = f(t), \quad t \in [a, b],$$

es decir, la sucesión de las sumas parciales

$$s_n(t) = \sum_{k=1}^{n} f_k(t)$$

de esta serie converge a la función f en el sentido de la media cuadrática:

$$\lim_{n\to\infty}\int_a^b [f(t) - s_n(t)]^2 dt = 0.$$

Entonces, de acuerdo con el lema 12, para toda función $\varphi(x) \in RL_2[a, b]$ tenemos

$$(f, \varphi) = \sum_{n=1}^{\infty}(f_n, \varphi),$$

es decir,

$$\int_a^b f(x)\varphi(x)dx = \sum_{n=1}^{\infty}\int_a^b f_n(x)\varphi(x)dx.$$

En particular, cuando $\varphi = 1$,

$$\int_a^b f(x)dx = \sum_{n=1}^{\infty}\int_a^b f_n(x)dx.$$

En otras palabras,

$$\int_a^b\left[\sum_{n=1}^{\infty} f_n(x)\right]dx = \sum_{n=1}^{\infty}\int_a^b f_n(x)dx.$$

Así pues, *si una serie de funciones de cuadrado integrable en el segmento* $[a, b]$ *converge en éste en el sentido de la media cuadrática hacia cierta función, también*

de cuadrado integrable en $[a, b]$, *entonces la serie puede integrarse término a término.*

Por cuanto de la convergencia uniforme de una sucesión de funciones continuas se deduce la convergencia de dicha sucesión a la misma función también en el sentido de la media cuadrática (véase el p. 57.4), de la afirmación aquí demostrada proviene que

si la serie de funciones continuas converge en un segmento uniformemente, se puede integrarla término a término.

Este mismo resultado se ha obtenido por otro procedimiento antes, en el capítulo sobre las series (véase el teorema 9 en el p. 36.4).

Definición 36. *Dos espacios lineales* X *e* Y *con producto escalar (semiescalar) se llaman isomorfos, si son isomorfos como espacios lineales, y la aplicación* f, *que aplica el espacio* X *sobre el espacio* Y *y realiza este isomorfismo, conserva el producto escalar (producto semiescalar), es decir, para cualesquiera dos elementos* $x \in X$ *e* $y \in X$ *se verifica la igualdad*

$$(x, y) = (f(x), f(y)).$$

Dos espacios lineales isomorfos provistos de producto escalar (semiescalar) pueden diferenciarse sólo en la naturaleza de sus elementos y no en las propiedades métricas, razón por la cual en adelante muy a menudo los espacios lineales isomorfos provistos de producto escalar (semiescalar) no se destinguirán.

Aclaremos esto con un ejemplo. Sean X e Y^* unos espacios lineales provistos de producto escalar (semiescalar) y sea f la aplicación isomorfa del espacio X en un conjunto $Y \subset Y^*$. En este caso, "al identificar" los elementos del espacio X con los elementos correspondientes del conjunto Y, podemos considerar el espacio X como un subespacio del espacio Y^*. Se entiende (compárese con las construcciones correspondientes en el p. 57.1 y p. 57.5) el estudio del espacio lineal X^* compuesto de los elementos del espacio X y los del conjunto $Y^* \setminus Y$. Las operaciones de sumación de los elementos y su multiplicación por un número en el espacio X^* se introducen igual que en el p. 57.5 (después de la definición 27), mientras que el producto escalar (semiescalar) $(x, y)_{X^*}$, $x \in X^*$, $y \in X^*$, se determina en el espacio a través del producto escalar (semiescalar) en el espacio Y^* mediante la biyección $F: X^* \to Y^*$, que viene definida por la fórmula (57.19) de la manera siguiente:

$$(x, y)_{X^*} = (F(x), F(y)),$$

donde en el segundo miembro figura el producto escalar (semiescalar) en el espacio Y^*. Es fácil comprobar que el espacio X^* es isomorfo al espacio Y^*.

Ejercicios. 22. Demuéstrese que todos los espacios lineales n-dimensionales provistos de producto escalar son isomorfos entre sí.

22. Demuéstrese que todo espacio lineal n-dimensional provisto de producto escalar es completo en el sentido de la métrica engendrada por el producto escalar.

Definición 37. *Un espacio lineal provisto de producto escalar, completo en el sentido de la métrica engendrada por el producto escalar dado, se llama espacio de Hilbert.* *)

*) D. Hilbert (1862 — 1943), matemático alemán.

Un espacio simplemente lineal provisto de producto escalar lo llaman también *prehilbertiano*. Esta denominación se justifica por el siguiente teorema.

Teorema 4. *Todo espacio prehilbertiano X se contiene y, además, es denso, en cierto espacio de Hilbert X^*.*

DEMOSTRACIÓN. De acuerdo con los teoremas 1 del p. 57.1 y 3 del p. 57.6, basta mostrar que a la completación X^* del espacio lineal normalizado X puede prolongarse desde X un producto escalar, conservando las propiedades $1^\circ \div 4^\circ$. Esto se puede hacer pasando al límite. En efecto, por cuanto $\overline{X} = X^*$, para cualquier par de puntos $x \in X^*$ e $y \in X^*$ existen las sucesiones de puntos $x_n \in X$, $y_n \in X$, $n = 1, 2, \ldots$, de tal género que

$$\lim_{n \to \infty} x_n = x, \quad \lim_{n \to \infty} y_n = y.$$

Probemos que existe $\lim_{n \to \infty} (x_n, y_n)$. Efectivamente, de la desigualdad (57.33) se infiere que para cualesquiera m y n naturales

$$|(x_m, y_m) - (x_n, y_n)| \leqslant \|x_m - x_n\| \|y_m\| + \|x_n\| \|y_m - y_n\|.$$

Puesto que las sucesiones $\{x_n\}$ e $\{y_n\}$, siendo convergentes, están acotadas en norma y son fundamentales, de la desigualdad mencionada proviene que la sucesión numérica $\{(x_n, y_n)\}$ es también fundamental y, por consiguiente, converge.

Pongamos, por definición, $(x, y) = \lim_{n \to \infty} (x_n, y_n)$. Es fácil comprobar, pasando al límite, que esta definición no depende de la elección de tales sucesiones $\{x_n\}$ e $\{y_n\}$ que $x_n \to x$, $y_n \to y$, y que para la función (x, y), definida de este modo se cumplen las propiedades $1^\circ - 4^\circ$ del producto escalar. □

El espacio obtenido de Hilbert se llama *completación del espacio prehilbertiano de partida.*

Como ejemplo de espacio de Hilbert interviene el espacio euclídeo n-dimensional (véase (57.8)). Otros ejemplos serán considerados más abajo.

Ejercicio 23. Demuéstrese que un espacio prehilbertiano, isomorfo al espacio de Hilbert, es un espacio de Hilbert.

57.10. ESPACIO L_2

Recordemos (véase el ejemplo 3 en el p. 57.8) que un espacio lineal de funciones continuas en el segmento $[a, b]$ provisto de producto escalar, definido según la fórmula (57.30), se denota mediante $CL_2[a, b]$.

La norma en el espacio $CL_2[a, b]$ se determina mediante la fórmula (57.31).

Lema 13. *Es espacio $CL_2[a, b]$ no es de Hilbert.*

DEMOSTRACIÓN. Para convencerse de que todo espacio $CL_2[a, b]$ no es completo, basta considerar el espacio $CL_2[a, b]$ para cierto segmento fijo (¿por qué?). Elijamos, para concretar, el segmento $[-1, 1]$ y demos un ejemplo de la sucesión de funciones fundamental en el espacio $CL_2[-1, 1]$ que no converge en dicho espacio.

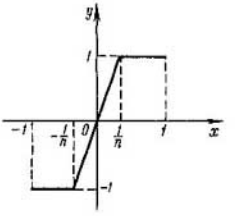


Fig. 237

Pongamos

$$f_n(x) = \begin{cases} -1, & \text{si} \quad -1 \leqslant x \leqslant -\dfrac{1}{n}, \\ nx, & \text{si} \quad -\dfrac{1}{n} < x < \dfrac{1}{n}, \\ 1, & \text{si} \quad \dfrac{1}{n} \leqslant x \leqslant 1, \end{cases} \tag{57.34}$$

$$n = 1, 2, \ldots$$

(fig. 237). Es evidente que las funciones $f_n(x)$, $n = 1, 2, \ldots$, son continuas en el segmento $[-1, 1]$. Observando luego que $|f_n(x)| \leqslant 1$, tenemos para $m > n$

$$\|f_n - f_m\|^2 = \int_{-1}^{1} |f_n(x) - f_m(x)|^2 dx = \int_{-1/n}^{1/n} |f_n(x) - f_m(x)|^2 dx \leqslant$$

$$\leqslant \int_{-1/n}^{1/n} [|f_n(x)| + |f_m(x)|]^2 dx \leqslant 4 \int_{-1/n}^{1/n} dx = \frac{8}{n}$$

de donde, evidentemente, se deduce que la sucesión (57.34) es fundamental en el espacio $CL_2[a, b]$.

Efectivamente, si está dado $\varepsilon > 0$, entonces, al elegir n_0 de modo tal que sea $8/n_0 < \varepsilon$, para todos los $n \geqslant n_0$ y cualesquiera $m > n$ tendremos $\|f_n - f_m\| < \dfrac{8}{n} \leqslant \dfrac{8}{n_0} < \varepsilon$. Como

$$\lim_{n \to \infty} f_n(x) = \begin{cases} -1 & \text{para} \quad -1 \leqslant x < 0, \\ 0 & \text{para} \quad x = 0, \\ 1 & \text{para} \quad 0 < x \leqslant 1, \end{cases}$$

es natural esperar que si la sucesión $[f_n]$ converge en el sentido de la media cuadrática, converge hacia la misma función, a la que converge puntualmente, es decir, hacia la función (véase fig. 238):

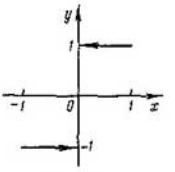


Fig. 238

$$f(x) \stackrel{\text{def}}{=} \begin{cases} -1 & \text{para} \quad -1 \leqslant x < 0, \\ 0 & \text{para} \quad x = 0 \\ 1 & \text{para} \quad 0 < x \leqslant 1. \end{cases}$$

Sin embargo, esta función f es discontinua, por lo cual $f \notin CL_2[0, 1]$. Por consiguiente, es natural esperar que la sucesión $\{f_n\}$ no tiene límite en el espacio $CL_2[a, b]$. Probémoslo.

No es difícil de convencerse de que la sucesión (57.34) converge en el segmento $[-1, 1]$ en el sentido de la seminorma (57.31) hacia la función f. En efecto,

$$\|f - f_n\|^{2*)} = \int_{-1}^{1} |f(x) - f_n(x)|^2 dx =$$

$$= \int_{-1/n}^{1/n} |f(x) - f_n(x)|^2 dx \leqslant \int_{-1/n}^{+1/n} [|f(x)| + |f_n(x)|^2] dx \leqslant$$

$$\leqslant 4 \int_{-1/n}^{+1/n} dx = \frac{8}{n} \to 0 \quad \text{para} \quad n \to \infty,$$

pues, $|f(x)| \leqslant 1$, $|f_n(x)| \leqslant 1$, $x \in [-1, 1]$. (57.35)

El límite según la seminorma no es único y por eso surge la cuestión: ¿existe o no, además, una función continua que también sirva de límite de la sucesión $\{f_n\}$ en el sentido de la media cuadrática? Mostremos que tal función no existe. Admitamos lo contrario. Supongamos que existe una función $g(x)$, continua en el segmento $[-1, 1]$, tal que

$$\lim_{n \to \infty} \|g - f_n\| = 0. \tag{57.36}$$

En este caso

$$\|f - g\| = \|(f - f_n) + (f_n - g)\| \leqslant \|f - f_n\| + \|f_n - g\|,$$

*) Por cuanto $f - f_n$ ya no es una función continua, el símbolo $\|\varphi\|$ significa aquí la seminorma (57.31) de la función φ. Esto debe tomarse en consideración también en los razonamientos que siguen.

donde ambos sumandos del segundo miembro, en virtud de (57.35) y (57.36), tienden a cero cuando $n \to \infty$, mientras que el primer miembro no depende de n, por consiguiente,

$$\int_{-1}^{1} |f(x) - g(x)|^2 dx = \|f - g\|^2 = 0;$$

y, con mayor razón,

$$\int_{-1}^{0} |f(x) - g(x)|^2 dx = 0, \quad \int_{0}^{1} |f(x) - g(x)|^2 dx = 0. \tag{57.37}$$

Examinemos, por ejemplo, el caso de $x \geqslant 0$. Dado que las funciones f y g son continuas en el intervalo $(0, 1)$, entonces, en virtud de (57.37), coinciden en dicho intervalo (véase la propiedad 9 de la integral definida en el p. 28.1). Por eso

$$g(+0) = \lim_{x \to +0} g(x) = \lim_{x \to +0} f(x) = 1.$$

Por analogía, considerando el caso cuando $x \leqslant 0$, tendremos

$$g(-0) = \lim_{x \to -0} f(x) = -1,$$

es decir, g es una función discontinua.

La contradición obtenida demuestra precisamente la afirmación. □

Así pues, el espacio lineal $CL_2[a, b]$ no es completo. No obstante, sabemos que todo espacio prehilbertiano puede ser complementado hasta que se obtenga un espacio completo, en particular, esto es cierto con relación al espacio en consideración. Volveremos a este problema un poco más abajo y ahora mismo analizaremos un espacio más.

Trataremos de abarcar una clase de funciones más amplia que las funciones continuas, a saber, consideraremos el espacio lineal $RL_2[a, b]$ de las funciones de cuadrado integrable en cierto segmento $[a, b]$ (véase el ejemplo 3 en el p. 57.8) y provisto de producto semiescalar que viene definido mediante la fórmula (57.30) y construyamos, valiéndonos de este espacio, otro espacio dotado de producto escalar.

Definición 38. *Dos funciones f y g de cuadrado integrable en el segmento $[a, b]$ las llamaremos equivalentes, si la seminorma* (57.31) *de su diferencia es igual a cero*:

$$\|f - g\| = \sqrt{\int_a^b [f(x) - g(x)]^2 dx} = 0. \tag{57.38}$$

La equivalencia de las funciones en el sentido de esta definición se designará por el símbolo.

$$f \sim g. \tag{57.39}$$

El empleo en este caso del mismo símbolo que se ha usado para designar la igualdad asintótica de las funciones, es decir, designar su equivalencia en el sentido del orden de su variación (véase la definición 3 en el p. 8.2), no nos llevará a equivocaciones algunas, puesto que cada vez quedará claro de qué equivalencia entre las funciones se trata en el caso dado.

La relación de equivalencia (57.39) posee las siguientes propiedades:

1°) $f \sim f$;

2°) si $f \sim g$, también $g \sim f$;

3°) si $f \sim g$ y $g \sim h$, entonces $f \sim h$.

Un conjunto de todas las funciones de cuadrado integrable en el segmento $[a, b]$, es decir, el espacio $RL_2[a, b]$ lo dividiremos en clases de las funciones equivalentes entre sí. Dichas clases se llamarán clases de equivalencia y se denotarán por letras latinas mayúsculas $F, G, H, \ldots$, y la totalidad de ellas, por $\mathfrak{F}$. Toda función f, perteneciente a la clase de equivalencia F, se denominará representante de dicha clase. Expresando brevemente el proceso de construcción del conjunto $\mathfrak{F}$, suele decirse que dicho conjunto se obtiene del conjunto de todas las funciones de cuadrado integrable por "identificación" de sus elementos equivalentes. Así pues, ahora cada conjunto de las funciones equivalentes se considera como un único elemento del conjunto $\mathfrak{F}$.

Para todo $F \in \mathfrak{F}$ y cada número real λ el elemento λF se determina de la manera siguiente. Elijamos un representante $f \in F$, entonces la función λf será también una función de cuadrado integrable en el segmento $[a, b]$, y por lo tanto, pertenece a cierta clase de equivalencia, es decir, será un representante de cierto elemento, perteneciente a $\mathfrak{F}$, el cual se determina precisamente como elemento λF.

Con el fin de mostrar que esta definición es correcta, conviene demostrar que el elemento λF no depende de cómo se elige la función $f \in F$. En efecto, si $f \in F$ y $f_1 \in F$, entonces $f_1 \sim f$, es decir, $\|f_1 - f\| = 0$. Por consiguiente, $\|\lambda f_1 - \lambda f\| = |\lambda| \|f_1 - f\| = 0$, lo que quiere decir que $\lambda f_1 \sim \lambda f$. I or esta razón las funciones λf_1 y λf pertenecen a una misma clase de equivalencia, es decir, a un mismo elemento del conjunto $\mathfrak{F}$.

Definamos ahora la operación de adición de los elementos del conjunto $\mathfrak{F}$. Sea $F \in \mathfrak{F}$ y $G \in \mathfrak{F}$. Elijamos unas funciones $f \in g$ y $f \in G$. Un elemento $F + G$ lo definamos como una clase de equivalencia que contiene el elemento $f + g$. Esta definición es unívoca, pues, si

$$f \in F, \quad f_1 \in F, \quad g \in G, \quad g_1 \in G$$

y, por consiguiente,

$$f_1 \sim f, \quad g_1 \sim g,$$

entonces

$$\|f_1 - f\| = 0, \quad \|g_1 - g\| = 0.$$

Por esto

$$0 \leqslant \|(f_1 + g_1) - (f + g)\| \leqslant \|f_1 - f\| + \|g_1 - g\| = 0,$$

es decir,

$$f_1 + g_1 \sim f + g$$

y, de este modo, la función $f_1 + g_1$ pertenece a la misma clase de equivalencia que la función $f + g$.

Así pues, para sumar los elementos del conjunto $\mathfrak{F}$ o multiplicarlos por un número, se deben elegir sus representantes, sobre los cuales ha de realizarse la operación indicada; de resultas se obtendrá cierta función; la clase de equivalencia, de cuyo representante es la citada función, será precisamente el resultado de la operación que se considera.

El conjunto $\mathfrak{F}$ con las operaciones introducidas λF y $F + G$ forma un espacio lineal. En efecto, para dichas operaciones se cumplen las propiedades 1°, 2°, 3° de la definición 11 en el p. 57.2. Comprobemos, por ejemplo, que para cualesquiera $F \in \mathfrak{F}$, $G \in \mathfrak{F}$ y todo número λ se verifica la desigualdad

$$\lambda(F + G) = \lambda F + \lambda G. \tag{57.40}$$

Si $f \in F$ y $g \in G$, entonces, de acuerdo con la definición de adición de elementos del conjunto $\mathfrak{F}$, obtendremos $f + g \in F + G$, $\lambda(f + g) \in \lambda(F + G)$. Por cuanto f y g son los elementos del espacio lineal, entonces $\lambda(f + g) = \lambda f + \lambda g$. En virtud de la regla de multiplicación de los elementos de $\mathfrak{F}$ por un número y de adición de estos elementos

$$\lambda f \in \lambda F, \quad \lambda g \in \lambda G, \quad \lambda f + \lambda g \in \lambda F + \lambda G.$$

De este modo, las clases de equivalencia $\lambda(F + G)$ y $\lambda F + \lambda G$ contienen un elemento común $\lambda(f + g) = \lambda f + \lambda g$, y, por lo tanto, coinciden. La igualdad (57.40) queda demostrada.

Análogamente se comprueba también el cumplimiento de las demás propiedades de los espacios lineales (véase la definición 11 en el p. 57.2) para las operaciones de adición y multiplicación por un número de los elementos pertenecientes al conjunto $\mathfrak{F}$.

Indiquemos que la clase de equivalencia que contiene una función idénticamente igual a cero en el segmento $[a, b]$ representa el cero del espacio lineal obtenido $\mathfrak{F}$. Esta clase se compone de aquellas, y sólo aquellas, funciones f que son equivalentes a cero, en otras palabras, para las cuales la seminorma (57.31) es igual a cero: $\|f\| = 0$, es decir,

$$\int_a^b f^2(x)dx = 0.$$

Definamos ahora en el espacio lineal $\mathfrak{F}$ la multiplicación escalar. Sea $F \in \mathfrak{F}$, $G \in \mathfrak{F}$; elijamos, en las clases F y G, ciertas representantes $f \in F$ y $g \in G$ y pongamos

$$(F, G) \stackrel{\text{def}}{=} (f, g). \tag{57.41}$$

De este modo, para multiplicar escalarmente los elementos del espacio $\mathfrak{F}$, es preciso elegir sus representantes y multiplicarlos escalarmente uno por otro (en el sentido del producto semiescalar (57.30)). El resultado obtenido será precisamente igual al producto escalar de los elementos en consideración del conjunto $\mathfrak{F}$.

La definición (57.41) tampoco depende de la elección de las funciones en las clases de equivalencia. En efecto, si

$$f \in F, \quad f_1 \in F, \quad g \in G, \quad g_1 \subset G,$$

entonces

$$f_1 \sim f, \quad g_1 \sim g$$

y, por consiguiente,

$$\|f_1 - f\| = 0, \|g_1 - g\| = 0.$$

Por eso, haciendo uso de la desigualdad de Cauchy — Schwarz (57.28), obtendremos

$$0 \leqslant |(f_1, g_1) - (f, g)| = |[(f_1, g_1) - (f, g_1)] + [(f, g_1) - (f, g)]| \leqslant$$
$$\leqslant |(f_1 - f, g_1)| + |(f_1, g_1 - g)| \leqslant \|f_1 - f\| \|g_1\| + \|f\| \|g_1 - g\| = 0.$$

De este modo $(f_1, g_1) = (f, g)$.

La función (57.41) satisface todas las propiedades de la multiplicación escalar. Efectivamente, sea $f \in F \in \mathfrak{F}$, $g \in G \in \mathfrak{F}$, $h \in H \in \mathfrak{F}$, y sean λ y μ unos números, entonces

$$(\lambda F + \mu G, H) = (\lambda f + \mu g, h) = \lambda(f, h) + \mu(g, h) = \lambda(F, H) + \mu(G, H),$$
$$(F, G) = (f, g) = (g, f) = (G, F),$$
$$(F, F) = (f, f) \geqslant 0.$$

Por fin, si $(F, F) = 0$, esto es testimonio de que para cualquier función $f \in F$ se tiene $(f, f) = \|f\|^2 = 0$, es decir, $f \sim 0$, que, de acuerdo con lo dicho anteriormente, significa precisamente que el elemento F es el elemento nulo del espacio $\mathfrak{F}$.

Definición 39. *El espacio lineal $\mathfrak{F}$ provisto de producto escalar* (57.41) *lleva el nombre de espacio $\widetilde{RL}_2 = \widetilde{RL}_2[a, b]$.*

Indiquemos que la norma $\|F\|_{\widetilde{RL}_2}$ del elemento F en el espacio $\widetilde{RL}_2[a, b]$ se determina, de conformidad con (57.27) y (57.41), en términos de la seminorma $\|f\|_{RL_2}$ de la función $f \in F$ según la fórmula

$$\|F\|_{\widetilde{RL}_2} = \|f\|_{RL_2} = \left[\int_a^b f^2(x)\,dx\right]^{1/2}, \quad f \in F, \tag{57.42}$$

con la particularidad de que, en virtud de la univocidad demostrada de la definición de producto escalar, la definición 39 es unívoca, es decir, no depende de la elección de la función $f \in F$.

OBSERVACIÓN 1. Los elementos del espacio $\widetilde{RL}_2[a, b]$ están constituidos por las clases de funciones equivalentes; no obstante, en la literatura matemática se encuentra a menudo la expresión "función del espacio $\widetilde{RL}_2$". Esta expresión convencional significa simplemente que se trata de una función de cuadrado integrable la que, por consiguiente, pertenece a una de las clases de funciones equivalentes que se consideran, es decir, es el representante de esta clase. Dicha expresión es cómoda, puesto que las operaciones de sumación, de multiplicación por un número y de multiplicación escalar de las clases de funciones equivalentes se reducen a la operación correspondiente sobre sus representantes, con la particularidad de que el resultado no depende de la elección de los representantes mencionados. Esta circunstancia justifica en cierto sentido una expresión de uso frecuente "el espacio $\widetilde{RL}_2[a, b]$ se compone de las funciones de cuadrado integrable", en este caso el espacio $\widetilde{RL}_2$ se designa a menudo simplemente mediante RL_2.

Toda función continua en el segmento $[a, b]$, siendo función de cuadrado integrable en este segmento, pertenece a cierta clase de equivalencia, es decir, a cierto elemento del espacio $\widetilde{RL}_2[a, b]$. Además, en la clase indicada no hay otra función continua, pues si las funciones continuas son equivalentes, son también iguales.

Estudiaremos una aplicación que a toda función continua $f \in CL_2[a, b]$ le pone en correspondencia la clase de equivalencia $F \in \widetilde{RL}_2[a, b]$, a la que esta función pertenece: $f \in F$. Esta aplicación se denomina *aplicación natural* de $CL_2[a, b]$ en $\widetilde{RL}_2[a, b]$. En virtud de la propia definición de las operaciones de sumación de elementos (que son las clases de equivalencia), de su multiplicación por un número y su producto escalar en el espacio $\widetilde{RL}_2[a, b]$ que se reducen a las mismas operaciones sobre los representantes de las clases de equivalencia, la aplicación natural es lineal y conserva el producto escalar. Es la aplicación biunívoca (inyección) del espacio $CL_2[a, b]$ en el espacio $\widetilde{RL}_2[a, b]$, puesto que si, realizándose esta aplicación, dos funciones continuas se aplicaran en un mismo elemento del espacio $\widetilde{RL}_2[a, b]$, es decir, en la misma clase de equivalencia, entonces ambas pertenecerían a esta clase. Lo último es posible, de acuerdo con lo afirmado más arriba, sólo en el caso en que representan una misma función continua.

Con el fin de estudiar las propiedades restantes de la aplicación natural, demostremos tres lemas sobre la aproximación de las funciones. Para abreviar, en lugar de $\|\cdot\|_{RL_2}$ escribiremos en estos lemas simplemente $\|\cdot\|$.

Lema 14. *Supongamos que el cuadrado de la función f es integrable en un intervalo, finito o infinito, cuyos extremos son a y b, $-\infty \leqslant a < b \leqslant +\infty$. En este caso para cualquier $\varepsilon > 0$ existe tal función φ* (véase el p. 55.2), *finita, escalonada e igual a cero fuera del intervalo mencionado, que*

$$\|f - \varphi\| < \varepsilon.$$

DEMOSTRACIÓN. Supongamos, para simplificar, que la función f es integrable según Riemann en cualquier segmento $[\xi, \eta]$, $a < \xi < \eta < b$, es decir, dentro del segmento en consideración con los extremos a y b no hay puntos singulares de la función f (véase el p. 55.1). El caso general se reduce fácilmente a éste.

Sea dado $\varepsilon > 0$. Fijemos ξ y η de modo tal que sea

$$\int_a^{\xi} f^2(x)dx + \int_{\eta}^{b} f^2(x)dx < \frac{\varepsilon^2}{2}. \tag{57.43}$$

Esto es posible, puesto que la integral extendida al segmento $[a, b]$ de la función f^2 es convergente. La función f, siendo integrable según Riemann en el segmento $[\xi, \eta]$, está acotada en él:

$$|f(x)| \leqslant M, \xi \leqslant x \leqslant \eta, \tag{57.44}$$

donde M es una constante.

De acuerdo con el lema 2 del p. 55.2, para $\varepsilon > 0$ dado existe una función escalonada finita φ tal que su portador supp φ está contenido en el segmento $[\xi, \eta]$, es decir, supp $\varphi \subset [\xi, \eta]$,

$$|\varphi(x)| \leqslant M, \quad x \in [\xi, \eta] \tag{57.45}$$

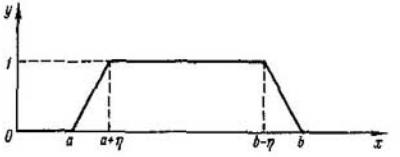


Fig. 239

(esto proviene de la fórmula (55.9)) y

$$\int_{\xi}^{\eta} |f(x) - \varphi(x)|\,dx < \frac{\varepsilon^2}{4M}. \tag{57.46}$$

Al aplicar sucesivamente las desigualdades (57.43), (57.44), (57.45) y (57.46), obtendremos:

$$\|f - \varphi\|^2 = \int_a^b [f(x) - \varphi(x)]^2 dx = \int_a^{\xi} f^2(x)dx + \int_{\xi}^{b} [f^2(x)dx +$$

$$+ \int_{\xi}^{\eta} [f(x) - \varphi(x)]^2 dx \leqslant \frac{\varepsilon^2}{2} + \int_{\xi}^{\eta} [|f(x)| + |\varphi(x)|]\,|f(x) - \varphi(x)|\,dx \leqslant$$

$$\leqslant \frac{\varepsilon^2}{2} + 2M \int_{\xi}^{\eta} |f(x) - \varphi(x)|\,dx < \frac{\varepsilon^2}{2} + 2M\,\frac{\varepsilon^2}{4M} = \varepsilon^2.$$

De aquí se deduce que $\|f - \varphi\| < \varepsilon$. □

Lema 15. *Sea φ una función escalonada finita, igual a cero fuera del segmento $[a, b]$; entonces, para cualquier $\varepsilon > 0$ existe una función g, finita, continua en todo el eje numérico e igual a cero fuera del segmento citado, de tal índole que*

$$\|g - \varphi\| < \varepsilon.$$

DEMOSTRACIÓN. Será suficiente considerar el caso de la función característica del semiintervalo, pues toda función escalonada finita es una combinación lineal finita de las funciones semejantes (véase el p. 55.2). Así pues, sea dada una función

$$\chi(x) = \begin{cases} 1 & \text{para } a \leqslant x < b, \\ 0 & \text{para } x < a \text{ y } x \geqslant b, \end{cases}$$

y dado también $\varepsilon > 0$. Tomemos un $\eta > 0$ de modo tal que se cumplan las desigualdades

$$\eta < \frac{\varepsilon^2}{8}, \quad \eta < \frac{b - a}{2},$$

y consideremos la función $g(x)$ cuya gráfica está expuesta en la fig. 239.

Dicha función puede ser escrita analíticamente del modo siguiente:

$$g(x) = \begin{cases} 0 & \text{para } x < a \text{ y } x > b, \\ \dfrac{x-a}{\eta} & \text{para } a \leqslant x \leqslant a+\eta, \\ 1 & \text{para } a+\eta < x < b-\eta, \\ \dfrac{b-x}{\eta} & \text{para } b-\eta \leqslant x \leqslant b. \end{cases}$$

Es evidente que $g(x)$ es una función finita continua en todo el eje numérico. Por cuanto $|\chi(x)| \leqslant 1$, $|g(x)| \leqslant 1$, $-\infty < x < +\infty$, entonces

$$\|\chi - g\|^2 = \int_a^b [\chi(x) - g(x)]^2 dx = \int_a^{a+\eta} [\chi(x) - g(x)]^2 dx +$$

$$+ \int_{b-\eta}^b [\chi(x) - g(x)]^2 dx \leqslant \int_a^{a+\eta} [|\chi(x)| + |g(x)|]^2 dx +$$

$$+ \int_{b-\eta}^b [|\chi(x)| + |g(x)|]^2 dx \leqslant 4 \int_a^{a+\eta} dx + 4 \int_{b-\eta}^b dx < 8\eta < \varepsilon^2,$$

es decir, $\|\chi - g\| < \varepsilon$. □

Lema 16. *Si f es una función de cuadrado intergable en el segmento $[a, b]$, constituye en este segmento el límite, en el sentido de la media cuadrática, para la sucesión de funciones finitas y continuas en todo el eje numérico f_n, $n = 1, 2, \ldots$, cuyos portadores se disponen en el segmento $[a, b]$:*

$$\lim_{n \to \infty} \|f_n - f\|^2 = \lim_{n \to \infty} \int_a^b [f_n(x) - f(x)]^2 dx = 0. \tag{57.47}$$

DEMOSTRACIÓN. Cualquiera que sea $\varepsilon > 0$, en virtud del lema 14, existe tal función finita escalonada φ, igual a cero fuera del segmento $[a, b]$, que

$$\|f - \varphi\| < \frac{\varepsilon}{2},$$

y, en virtud del lema 15, para esta función escalonada φ se encontrará tal función g, continua en todo el eje numérico e igual a cero fuera del segmento $[a, b]$, que

$$\|\varphi - g\| < \frac{\varepsilon}{2},$$

y, por lo tanto, (fig. 240)

$$\|f - g\| \leqslant \|f - \varphi\| + \|\varphi - g\| < \varepsilon.$$

Al elegir ahora una sucesión numérica ε_n que tiende a $+0$ cuando $n \to \infty$, $n = 1, 2, \ldots$, y designar por f_n una función continua en todo el eje numérico que, en virtud de la construcción citada, corresponde al número ε_n y es igual a cero fuera del segmento $[a, b]$, obtendremos la sucesión buscada $[f_n]$ que satisface la condición (57.47) (la definición del límite de la sucesión de funciones en el sentido de la media

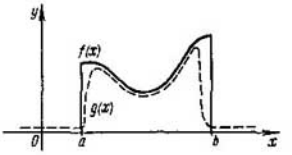


Fig. 240

cuadrática véase en el p. 57.5) y es tal que $\operatorname{supp} f_n \subset [a, b]$ para todo $n = 1, 2, \ldots$. □

Definición 40. *Un subconjunto del espacio $CL_2[a, b]$, compuesto de las funciones f que se reducen a cero en los extremos del segmento $[a, b]$: $f(a) = f(b) = 0$, se llama espacio $\mathring{C}L_2[a, b]$.*

Obviamente, el lema 16 significa que toda función de cuadrado integrable en el segmento $[a, b]$ puede aproximarse, con cualquier grado de precisión en el sentido de la media cuadrática, mediante las funciones pertenecientes a $\mathring{C}L_2[a, b]$. Está claro que $\mathring{C}L_2[a, b]$ es un espacio lineal prehilbertiano y

$$\mathring{C}L_2[a, b] \subset CL_2[a, b]. \tag{57.48}$$

Volveremos ahora a la aplicación natural

$$CL_2[a, b] \to \widetilde{RL}_2[a, b].$$

Teorema 5. *La aplicación natural $CL_2[a, b] \to \widetilde{RL}_2[a, b]$, es decir, aquella que a toda función continua en el segmento $[a, b]$ le pone en correspondencia una clase de equivalencia a la que dicha función pertenece, es la aplicación isomorfa del espacio $CL_2[a, b]$ en $\widetilde{RL}_2[a, b]$, con la particularidad de que la imagen del espacio, $\mathring{C}L_2[a, b]$ (y, por consiguiente, en virtud de (57.48), de todo el espacio $CL_2[a, b]$ también) es densa en $\widetilde{RL}_2[a, b]$.*

DEMOSTRACIÓN DEL TEOREMA 5. Designemos con Φ la aplicación natural del espacio $CL_2[a, b]$ en el espacio $\widetilde{RL}_2[a, b]$, es decir, aquella que a toda función f, continua en el segmento $[a, b]$, le pone en correspondencia una clase de funciones equivalentes de cuadrado integrable en dicho segmento a la que f pertenece, en otras palabras, la clase de equivalencia, de cuyo representante interviene la propia f. De este modo, si

$$f \in CL_2[a, b] \quad \text{y} \quad f \in F \in \widetilde{RL}_2[a, b],$$

entonces $\Phi(f) = F$.

Sea $F = \Phi(f) = 0$; en este caso $\|F\| = 0$, pero $f \in F$, por lo cual, también $\|f\| = 0$. De conformidad con la propiedad de la norma, de aquí se deduce que $f = 0$, es decir, el núcleo de la aplicación Φ se compone sólo del elemento nulo. Por cuanto la aplicación natural Φ es lineal, aplica biunívocamente el espacio $CL_2[a, b]$ en el espacio $\widetilde{RL}_2[a, b]$ (véase el lema 2 en el p. 57.2).

Mostremos que la *imagen* del espacio $\mathring{C}L_2[a, b]$ en esta aplicación es un conjunto denso en el espacio $\widetilde{RL}_2[a, b]$. Supongamos que $F \in \widetilde{RL}_2[a, b]$ y la función f es el

representante del segmento F, es decir, $f \in F$. Como f es una función de cuadrado integrable en el segmento $[a, b]$, entonces, de acuerdo con el lema 3, es el límite en el sentido de la media cuadrática para cierta sucesión de las funciones f_n, continuas en el segmento $[a, b]$, que se reducen a cero en los extremos de éste (véase (57.47)), es decir $f_n \in \mathring{C}L_2[a, b]$, $n = 1, 2, \ldots$. Si $f_n \in F_n \in \widetilde{RL}_2[a, b]$, entonces, según la definición de norma en el espacio $\widetilde{RL}_2[a, b]$, obtenemos

$$\|F_n - F\|_{\widetilde{RL}_2} = \|f_n - f\|_{RL_2},$$

donde a la derecha figura, como siempre, la seminorma (57.31). De aquí, en virtud de la igualdad (57.47), se obtiene

$$\lim_{n \to \infty} \|F_n - F\| = 0. \tag{57.49}$$

Por cuanto la clase de equivalencia F ha sido un elemento fijo arbitrario del espacio $\widetilde{RL}_2[a, b]$, y $F_n = \Phi(f_n)$, donde f_n es una función continua en el segmento $[a, b]$ que se reduce a cero en los extremos de este segmento y, por lo tanto, $F_n \in \Phi(\mathring{C}L_2[a, b])$, $n = 1, 2, \ldots$, entonces la igualdad (57.49) significa precisamente la densidad de la imagen del conjunto $\mathring{C}L_2[a, b]$ en el espacio $\widetilde{RL}_2[a, b]$ en el transcurso de la aplicación Φ.

Para demostrar la densidad de la imagen del conjunto $CL_2[a, b]$ en la aplicación natural de éste en el espacio $\widetilde{RL}_2[a, b]$, observemos que de la inclusión (57.48) se infiere, evidentemente, que

$$\Phi(CL_2[a, b]) \subset \Phi(CL_2[a, b]) \subset \widetilde{RL}_2[a, b].$$

Mientras tanto, si en algún espacio métrico X es denso el conjunto A, es decir, si $\overline{A} = X$ y $A \subset B \subset X$, entonces, por supuesto, el conjunto B es también denso en X, pues $A \subset \overline{B} \subset X$, y, como $\overline{A} = X$, se tiene también que $\overline{B} = X$. Por eso, de la densidad del conjunto $\Phi(\mathring{C}L_2[a, b])$ en el espacio $\widetilde{RL}_2[a, b]$ se desprende que el conjunto $\Phi(CL_2[a, b])$ es también denso en él. □

Identificando toda función continua $f \in CL_2[a, b]$ con la clase de funciones equivalentes $F \in \widetilde{RL}_2[a, b]$ a la que la función f pertenece: $f \in F$, es decir, identificando f con su imagen en la aplicación natural Φ, llegamos a que $CL_2[a, b]$ es un subconjunto del espacio $\widetilde{RL}_2[a, b]$:

$$CL_2[a, b] \subset \widetilde{RL}_2[a, b]. \tag{57.50}$$

Esta inclusión lleva el nombre de *encaje natural* del espacio CL_2 en el espacio $\widetilde{RL}_2$.

Así pues, en virtud de (57.48) y (57.50), resultan lícitas las inclusiones

$$\mathring{C}L_2[a, b] \subset CL_2[a, b] \subset \widetilde{RL}_2[a, b],$$

con la particularidad de que, de acuerdo con el teorema 5,

$$\overline{\mathring{C}L_2[a, b]} = \widetilde{RL}_2[a, b],$$

es decir, $\mathring{C}L_2[a, b]$, y, por ende, $CL_2[a, b]$, son densos en el espacio $\widetilde{RL}_2[a, b]$.

Se puede mostrar que el espacio $\widetilde{RL}_2[a, b]$ no es completo, es decir, no es espacio de Hilbert.

Problema 37. Demuéstrese que el espacio $\widetilde{RL}_2[a, b]$ no es completo.

En el p. 57.9 se ha demostrado que todo espacio prehilbertiano puede ser completado hasta que se obtenga un espacio completo, es decir, un espacio de Hilbert. En el caso particular, esto puede realizarse con relación al espacio $\widetilde{RL}_2[a, b]$.

Definición 41. *La completación del espacio prehilbertiano* $\widetilde{RL}_2 = \widetilde{RL}_2[a, b]$ *se denomina espacio* $L_2 = L_2[a, b]$.

Por definición de la completación tenemos

$$\widetilde{RL}_2[a, b] \subset L_2[a, b] \tag{57.51}$$

y $\widetilde{RL}_2[a, b]$ es denso en el espacio $L_2[a, b]$, es decir,

$$\overline{\widetilde{RL}_2[a, b]} = L_2[a, b].$$

En virtud de las inclusiones (57.48), (57.50) y (57.51) tienen lugar los encajes naturales

$$\mathring{CL}_2[a, b] \subset CL_2[a, b] \subset \widetilde{RL}_2[a, b] \subset L_2[a, b]. \tag{57.52}$$

Resulta que no sólo $\widetilde{RL}_2$ es denso en el espacio L_2, sino que $\mathring{CL}_2$ es también denso en L_2.

Teorema 6. *El espacio* $\mathring{CL}_2[a, b]$ *es denso en el espacio* $L_2[a, b]$.

Corolario. *El espacio* $CL_2[a, b]$ *es denso en el espacio* $L_2[a, b]$.

DEMOSTRACIÓN DEL TEOREMA 6. Sea $f \in L_2[a, b]$ y sea arbitrariamente fijado $\varepsilon > 0$. Para simplificar, todos los elementos del espacio $L_2[a, b]$ se designarán por letras latinas minúsculas, aunque dichos elementos no son, en el caso general, funciones. Por cuanto el espacio $L_2[a, b]$ es una completación del espacio $\widetilde{RL}_2[a, b]$, existe obligatoriamente tal elemento $g \in \widetilde{RL}_2[a, b]$ que

$$\|f - g\|_{L_2} < \frac{\varepsilon}{2}.$$

De conformidad con la inclusión (57.52) y densidad del conjunto $\mathring{CL}_2[a, b]$, en el espacio $\widetilde{RL}_2[a, b]$ existe tal elemento $h \in \mathring{CL}_2[a, b]$, que

$$\|g - h\|_{L_2} < \frac{\varepsilon}{2}.$$

Por eso

$$\|f - h\|_{L_2} \leqslant \|f - g\|_{L_2} + \|g - h\|_{L_2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto precisamente implica que el conjunto $\mathring{CL}_2[a, b]$ es denso en el espacio $L_2[a, b]$. □

El corolario se deduce naturalmente del teorema, puesto que (como se ha probado en la demostración del teorema 5) si un subconjunto A de cierto conjunto B, $A \subset B$, es denso en algún espacio métrico $X \supset B$, entonces el mismo conjunto B es, con mayor razón, denso en X. En el caso dado $\mathring{CL}_2[a, b] \subset CL_2[a, b] \subset L_2[a, b]$ y $\mathring{CL}_2[a, b]$ es denso en $L_2[a, b]$. Por eso, $CL_2[a, b]$ es también denso en $L_2[a, b]$. □

Ejercicio 24. Demuéstrese que si X es un espacio métrico, $A \subset B \subset X$, el conjunto A es denso en el conjunto B, y si el conjunto B es denso en el espacio X, entonces el conjunto A es también denso en el espacio X.

OBSERVACIÓN 2. Si consideramos $L_2[a, b]$ como un espacio que se obtiene del espacio $\widetilde{RL}_2[a, b]$ mediante la construcción (descrita en los teoremas 1, 3, 4 del presente párrafo) de la completación de espacios, entonces, como sus elementos intervendrán las clases de sucesiones fundamentales equivalentes compuestas de las clases de funciones equivalentes de cuadrado integrable. Si, en este caso, realizamos la identificación de los espacios CL_2 y RL_2 con sus imágenes en L_2, según se ha indicado más arriba, y de este modo consideramos que

$$CL_2 \subset \widetilde{RL}_2 \subset L_2,$$

resultará pues que el espacio L_2 queda compuesto de las funciones continuas, de las clases de funciones equivalentes de cuadrado integrable privadas de funciones continuas, y de los "elementos abstractos" que representan las clases indicadas de sucesiones fundamentales. Luego, todos los elementos del espacio $\widetilde{RL}_2$ pueden ser "sustituidos", de modo convencional en el sentido de la observación 1, por unas funciones, esto es, por sus representantes arbitrariamente elegidos. Entonces el espacio L_2 resultará compuesto de las funciones de cuadrado integrable y los mismos elementos abstractos que surgen forzosamente en el transcurso de completación del espacio $\widetilde{RL}_2$, en vista de que no es completo. Esta "sustitución convencional" de los elementos del espacio $\widetilde{RL}_2[a, b]$ por sus representantes refleja una afirmación exacta de que las operaciones sobre las clases de funciones equivalentes se reducen a las operaciones correspondientes con sus representantes en el sentido indicado arriba.

Resulta pues, y esto es muy interesante e importante, que los elementos abstractos mencionados pueden considerarse no como las clases de sucesiones fundamentales de las clases de equivalencia, sino como ciertas funciones, con más precisión, como clases de funciones equivalentes en el sentido de la definición 38, con la particularidad de que el producto escalar para ellas se determina también mediante las fórmulas (57.30) y (57.41) con la única peculiaridad consistente en que la integral en dichas fórmulas conviene entenderla no en el sentido de la integral de Riemann, propia o impropia, sino en el sentido de la así llamada integral de Lebesgue. El análisis de este problema sale, sin embargo, de los márgenes de los métodos que se consideran y no se realizará en el curso dado. Se puede encontrar dicho análisis en el excelente libro de texto "Curso del análisis matemático" de S. M. Nikolski.

OBSERVACIÓN 3. La definición del espacio $L_2[a, b]$ se generaliza también del modo natural al caso de un intervalo infinito. Consideraremos, para concretar, todo el eje numérico. Para dos funciones φ y ψ, continuas e integrables en cuadrado en todo el eje real, el producto escalar se determinará según la fórmula

$$(\varphi, \psi) = \int_{-\infty}^{+\infty} \varphi(x)\psi(x)dx. \tag{57.53}$$

Esta definición es correcta, pues la integral que figura en el segundo miembro converge, asumidas las suposiciones respecto de las funciones φ y ψ, y converge, además, absolutamente. Esto proviene inmediatamente de la desigualdad

$$|\varphi(x)\psi(x)| \leqslant \frac{\varphi^2(x) + \psi^2(x)}{2}.$$

Las propiedades del producto escalar para (57.53) se comprueban fácilmente. Se puede mostrar, por analogía con el caso de un intervalo finito, que el espacio métri-

co (que se obtiene en este caso) de funciones continuas e integrables en cuadrado, al igual que el espacio prehilbertiano que se obtiene mediante la "identificación" de las funciones equivalentes de cuadrado integrable en todo el eje numérico, no es completo en la métrica engendrada por el producto escalar (57.53). Las completaciones de estos espacios coinciden salvo un isomorfismo, y se designan con $L_2(-\infty, +\infty)$.

Ejercicios. 25. Demuéstrese que la función $f(x) = \frac{1}{\sqrt{x}}$ en el segmento [0, 1] no es un límite, en el sentido de la media cuadrática, de una sucesión de funciones continuas.

26. Demuéstrese que para una sucesión de funciones los conceptos de convergencia en media en el sentido de L_1 y en el sentido de L_2 no son equivalentes.

27. Demuéstrese que si una sucesión de funciones integrables en cierto segmento converge uniformemente en el mismo hacia cierta función integrable en él, entonces dicha sucesión en el segmento mencionado converge hacia la misma función y converge en media tanto en el sentido de L_1, como en el de L_2.

28. Constrúyase un ejemplo de sucesión de funciones, continuas en cierto segmento, que converja en él hacia cierta función continua en media en el sentido de L_2, pero no converja uniformemente en dicho segmento.

29. Constrúyase un ejemplo de sucesión de funciones continuas no negativas en un segmento que converja en él en media, pero no converja en el sentido de la media cuadrática.

Problema 38. Demuéstrese que para cualquier p, $1 \leqslant p < +\infty$, y cualquier intervalo con extremos en los puntos a y b, $-\infty \leqslant a < b \leqslant +\infty$, el conjunto de funciones continuas en este intervalo es denso en el espacio $\widetilde{RL}_p(a, b)$.

Hemos descrito diferentes tipos de los espacios. En el análisis matemático se estudian, principalmente, los espacios cuyos elementos están constituidos por funciones. Los espacios de este género se denominan *funcionales*.

En los ejemplos citados se consideraban, para simplificar, las funciones de una sola variable. De un modo similar, si tomamos un espacio lineal de funciones continuas en la clausura de cierto conjunto $G \subset R^n$, medible según Jordan, introducimos un producto escalar de acuerdo con la fórmula $(\varphi, \psi) = \int \varphi\psi dG$ y completamos a continuación el espacio obtenido, obtendremos un espacio de Hilbert que se denota con $L_2(G)$.

Se puede mostrar en este caso que todos los espacios $L_2(G)$, obtenidos de este modo, serán espacios de Hilbert separables de dimensión infinita.

El hecho de que el espacio $L_2[a, b]$ es de dimensión infinita será confirmado en el p. 58.2 y la separabilidad de este espacio se demostrará en el p. 58.3 (teorema 2).

En adelante (véase el p. 58.5, teorema 10) demostraremos que todos los espacios de Hilbert separables de dimensión infinita son isomorfos entre sí. De este modo, al estudiar las propiedades determinadas de las funciones de una o varias variables, se logra formar los espacios L_2 de algunos de los conjuntos de dichas funciones. No obstante, habiéndose convertido en los puntos de estos espacios, las funciones pierden muchas de sus propiedades individuales. En particular, los espacios L_2 no se diferencian uno del otro en lo que se refiere al número de variables, de las cuales dependen las funciones que forman parte de estos espacios. Naturalmente, esta circunstancia no impide de ninguna manera de que los espacios funcionales se empleen con gran éxito tanto en los problemas netamente teóricos, como en las aplicaciones de las matemáticas.

Las numerosas definiciones introducidas en este párrafo se utilizarán en lo que sigue para describir ciertas propiedades de diferentes clases de funciones en términos geométricos habituales y demostrativos (espacio, punto, distancia, vector, base, etc.) y ayudarán a establecer la analogía que existe entre los espacios vectoriales n-dimensionales ordinarios y los de funciones y aclarar, además, las propiedades específicas de los espacios funcionales de dimensión infinita.

§ 58. BASES ORTONORMALIZADAS Y DESARROLLOS SEGÚN ELLAS

58.1. SISTEMAS ORTONORMALIZADOS

Definición 1. *Sea X un espacio lineal con un producto semiescalar. Los elementos $x \in X$ e $y \in X$ se denominan ortogonales, si $(x, y) = 0$, y en este caso se anota también $x \perp y$.*

Definición 2. *Un sistema de elementos x_α, $\alpha \in \mathfrak{A}$ ($\mathfrak{A}$ es cierto conjunto de índices) del espacio lineal X provisto de producto semiescalar se denomina ortogonal, si dos elementos cualesquiera suyos son ortogonales. Si, además, la norma de todo elemento suyo es igual a la unidad, es decir, $\|x_\alpha\| = 1$, $\alpha \in \mathfrak{A}$, el sistema se llama ortonormalizado.*

Evidentemente, si el sistema x_α, $\alpha \in \mathfrak{A}$, es ortogonal y $x_\alpha \neq 0$, cualquiera que sea $\alpha \in \mathfrak{A}$, aquél se puede "normalizar". Efectivamente, al dividir todo elemento por su norma, es decir, al multiplicar x_α por el número $1/\|x_\alpha\|$, obtendremos un sistema ortonormalizado

$$\left\{ \frac{x_\alpha}{\|x_\alpha\|}, \alpha \in \mathfrak{A} \right\}.$$

Recordemos que si X es un espacio con un producto escalar, la condición $\|x\| \neq 0$ es equivalente a que $x \neq 0$.

Lema 1. *Si un sistema $\{x_\alpha, \alpha \in \mathfrak{A}\}$ de elementos del espacio X provisto de producto semiescalar es ortogonal y $\|x_\alpha\| \neq 0$, para cualquier $\alpha \in A$, dicho sistema es linealmente independiente.*

DEMOSTRACIÓN. Supongamos que para ciertos elementos

$$x_{\alpha_k}, \quad \alpha_k \in \mathfrak{A}, \quad k = 1, 2, \ldots, n,$$

se tiene

$$\lambda_1 x_{\alpha_1} + \lambda_2 x_{\alpha_2} + \ldots + \lambda_n x_{\alpha_n} = 0.$$

Multipliquemos escalarmente ambos miembros de esta igualdad por x_{α_k}, k está fijado ($k = 1, 2, \ldots, n$), obtendremos

$$\lambda_k (x_{\alpha_k}, x_{\alpha_k}) = 0,$$

pues, por ser ortogonal el sistema $(x_{\alpha_j}, x_{\alpha_k}) = 0$, $j \neq k$. Al observar luego que, por hipótesis, $x_{\alpha_k} \neq 0$, y, por lo tanto, $(x_{\alpha_k}, x_{\alpha_k}) \neq 0$, obtenemos $\lambda_k = 0$, $k = 1, 2, \ldots, n$.

La independencia lineal del sistema x_α, $\alpha \in \mathfrak{A}$, está demostrada. □

Demostremos un lema más que expresa el criterio de independencia lineal de las funciones a través de los productos escalares.

Lema 2. *Si para el sistema de elementos $x_1, \ldots, x_n$ del espacio X provisto de producto escalar el determinante*

$$G(x_1, \ldots, x_n) = \begin{vmatrix} (x_1, x_1)(x_1, x_2) \ldots (x_1, x_n) \\ (x_2, x_1)(x_2, x_2) \ldots (x_2, x_n) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ (x_n, x_1)(x_n, x_2) \ldots (x_n, x_n) \end{vmatrix}$$

es igual a cero, el sistema es linealmente dependiente.

El determinante $G(x_1, \ldots, x_n)$ recibe el nombre de *determinante de Gram**) para el sistema dado.

DEMOSTRACIÓN. Consideraremos un sistema de n ecuaciones lineales con n incógnitas λ_i, $i = 1, 2, \ldots, n$:

$$(\lambda_1 x_1 + \ldots + \lambda_n x_n, x_i) = 0, \quad i = 1, 2, \ldots, n, \tag{58.1}$$

o bien

$$\lambda_1 (x_1, x_i) + \ldots + \lambda_n (x_n, x_i) = 0, \quad i = 1, 2, \ldots, n.$$

El determinante de este sistema será un determinante transpuesto de Gram, el cual, por hipótesis del lema, es igual a cero. Por consiguiente, el sistema (58.1) tiene una solución no trivial $\lambda_1, \ldots, \lambda_n$ (es decir, una solución tal que no todas de las λ_i, $i = 1, 2, \ldots$, son nulas). Multipliquemos la igualdad (58.1) por λ_i y sumemos según i desde 1 hasta n:

$$(\lambda_1 x_1 + \ldots + \lambda_n x_n, \quad \lambda_1 x_1 + \ldots + \lambda_n x_n) = 0.$$

De aquí, $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$, lo que expresa la dependencia lineal del sistema $x_1, \ldots, x_n$. □

Ejercicios. 1. Demuéstrese que si un sistema finito de elementos de un espacio prehilbertiano es linealmente dependiente, su determinante de Gram es igual a cero.

2. Demuéstrese que si $\{\omega_\alpha\}$ es un sistema ortonormalizado, entonces para cualesquiera dos elementos suyos ω_α y $\omega_{\alpha'}$, tiene lugar la igualdad

$$\|\omega_{\alpha'} - \omega_\alpha\| = \sqrt{2}, \quad \alpha' \neq \alpha.$$

3. Demuéstrese que las funciones $\operatorname{sen} x$, $\operatorname{sen} 3x$, $\operatorname{sen} 5x$, $\operatorname{sen} 7x$, $\operatorname{sen} 9x$ son linealmente independientes sobre cualquier intervalo de longitud positiva.

Ejemplos. 1. El sistema trigonométrico de funciones 1,

$$1, \cos x, \operatorname{sen} x, \cos 2x, \operatorname{sen} 2x, \ldots, \cos nx, \operatorname{sen} nx, \ldots \tag{58.2}$$

es ortogonal en el espacio $L_2[-\pi, \pi]$ (véase el p. 57.10). Esto fue demostrado en el lema 1 del p. 55.1.

De las fórmulas (55.4) proviene que $\|\operatorname{sen} nx\| = \sqrt{\pi}$, $\|\cos nx\| = \sqrt{\pi}$, $n = 1$,

*) J. P. Gram (1850 — 1916), matemático danés.

2, . . . , por lo cual el sistema ortonormalizado, correspondiente al sistema (58.2), tiene por expresión

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\operatorname{sen} x, \ldots, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\operatorname{sen} nx \ldots .$$

2. Los polinomios

$$P_0(x) = 1, \quad P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2-1)^n}{dx^n}, \quad n = 1, 2, \ldots . \tag{58.3}$$

llevan el nombre de *Legendre*. De la fórmula (58.3) se ve que $P_n(x)$ es un polinomio de grado n:

$$P_n(x) = \frac{(2n-1)!!}{n!}x^n + \ldots$$

Mostremos que el sistema (58.3) es ortogonal en el espacio $L_2[-1, 1]$. Con este fin demostremos una afirmación más general, a saber, el polinomio de Legendre $P_n(x)$ es ortogonal a todo polinomio $Q_m(x)$ de grado $m < n$. Al notar previamente que la expresión

$$\frac{d^k(x^2-1)^n}{dx^k}$$

se anula, para $k = 0, 1, 2, \ldots, n-1$, en los puntos $x = -1$ y $x = 1$ tenemos, integrando por partes:

$$\int_{-1}^{1} Q_m(x)\frac{d^n(x^2-1)^n}{dx^n}dx = Q_m(x)\frac{d^{n-1}(x^2-1)^n}{dx^{n-1}}\Bigg|_{-1}^{1} -$$

$$-\int_{-1}^{1} Q_m'(x)\frac{d^{n-1}(x^2-1)^n}{dx^{n-1}}dx = \ldots = (-1)^m Q_m^{(m)}(x)\int_{-1}^{1}\frac{d^{n-m}(x^2-1)^n}{dx^{n-m}}dx =$$

$$= (-1)^m Q_m^{(m)}(x)\frac{d^{n-m-1}(x^2-1)^n}{dx^{n-m-1}}\Bigg|_{-1}^{1} = 0.$$

De este modo

$$\int_{-1}^{1} Q_m(x)P_n(x)dx = 0, \quad m < n;$$

en particular

$$\int_{-1}^{1} P_m(x)P_n(x)dx = 0, \quad m \neq n.$$

Calculemos ahora la norma de los polinomios de Legendre. Al observar que

$$P_n(x) = \frac{(2n-1)!!}{n!}x^n + Q_{n-1}(x),$$

donde $Q_{n-1}(x)$ es un polinomio de grado no superior a $n-1$, y hacer uso de la

ortogonalidad de $P_n(x)$ respecto de todos los polinomios de grado inferior, obtendremos:

$$\int_{-1}^{+1} P_n^2(x)dx = \int_{-1}^{+1} P_n(x)\left[\frac{(2n-1)!!}{n!}x^n + Q_{n-1}(x)\right]dx =$$

$$= \frac{(2n-1)!!}{n!}\int_{-1}^{1} P_n(x)x^n dx = \frac{(2n-1)!!}{n!(2n)!!}\int_{-1}^{1}\frac{d^n(x^2-1)^n}{dx^n}x^n dx.$$

Integrando sucesivamente por partes, tendremos

$$\int_{-1}^{+} P_n^2(x)dx = \ldots = (-1)^{n-1}\frac{(2n-1)!!}{(2n)!!}\int_{-1}^{1} xd(x^2-1)^n =$$

$$= (-1)^{n-1}\frac{(2n-1)!!}{(2n-2)!!}\int_{-1}^{1}(x^2-1)^{n-1}x^2dx =$$

$$= (-1)^{n-1}\frac{(2n-1)!!}{(2n-2)!!}\frac{1}{3}\int_{-1}^{1}(x^2-1)^{n-1}d(x^3) =$$

$$= (-1)^{n-1}\frac{(2n-1)!!}{(2n-4)!!}\frac{1}{3}\int_{-1}^{1}(x^2-1)^{n-2}x^4dx = \ldots$$

$$\ldots = \int_{-1}^{1} x^{2n}dx = \frac{2}{2n+1}.$$

De este modo,

$$\|P_n(x)\| = \sqrt{\frac{2}{2n+1}}.$$

El sistema de polinomios de Legendre, al igual que cualquier sistema ortogonal de elementos no nulos, es linealmente independiente (véase el lema 1) en el espacio $L_2[-1, 1]$. Por cuanto el sistema de funciones que se considera se compone aquí de los polinomios, la independencia lineal de éstos en un segmento (en el caso dado, en el segmento $[-1, 1]$) predetermina su independencia lineal en cualquier otro segmento.

En efecto, si ciertos polinomios $Q_1(x), \ldots, Q_k(x)$ son linealmente independientes en el segmento $[a, b]$, entonces, evidentemente, son linealmente independientes también en todo el eje numérico (todo sistema de funciones, linealmente independiente en cierto conjunto, es linealmente independiente también en cualquier conjunto mayor en el cual están definidas todas las funciones del sistema en consi-

deración). Si los polinomios $Q_1(x), \ldots, Q_k(x)$ resultaran ser linealmente dependientes en cierto segmento $[\alpha, \beta]$, es decir, si se encontraran tales números $\lambda_1, \ldots, \lambda_k$, no todos iguales a cero, que para cualquier $x \in [\alpha, \beta]$ se verificase la igualdad $\lambda_1 Q_1(x) + \ldots + \lambda_k Q_k(x) = 0$, esto significaría que todos los coeficientes del polinomio $\lambda_1 Q_1(x) + \ldots + \lambda_k Q_k(x)$ son nulos (un polinomio con los coeficientes distintos de cero puede tener sólo un número finito de ceros). Esto es indicio de que los polinomios $Q_1(x), \ldots, Q_k(x)$ son linealmente dependientes en todo el eje numérico. La contradicción obtenida demuestra su independencia lineal en el segmento $[\alpha, \beta]$.

De la independencia lineal de los polinomios de Legendre proviene que todo polinomio de grado no superior a n, es una combinación lineal de los polinomios de Legendre $P_0(x), P_1(x), \ldots, P_n(x)$. En efecto, en un espacio $(n+1)$-dimensional de los polinomios cuyos grados no sobrepasan n cualquier sistema de $n+1$ polinomios linealmente independientes, en particular, el sistema citado de polinomios de Legendre, forma una base. Por eso todo polinomio de grado mencionado es una combinación lineal de elementos del sistema indicado.

3. El sistema de funciones $\{e^{inx}\}$, $n = 0, \pm 1, \pm 2, \ldots$, es ortogonal en el segmento $[-\pi, \pi]$.

En efecto,

$$\int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}dx = \int_{-\pi}^{\pi} e^{i(n-m)x}dx.$$

De aquí, recordando que el período de la función e^z es igual a $2\pi i$ (véase el p. 37.6), para $n \neq m$ obtenemos:

$$\int_{-\pi}^{\pi} e^{inx}e^{imx}dx = \frac{1}{i(n-m)} e^{i(n-m)x}\Big|_{-\pi}^{\pi} = 0.$$

Ejercicio 4. Demuéstrese que la sucesión de funciones $\operatorname{sen}(2n-1)\frac{x}{2}$, $n = 1, 2, \ldots$, forma un sistema ortogonal en el segmento $[0, \pi]$.

58.2. ORTOGONALIZACIÓN

Supongamos nuevamente que X es un espacio prehilbertiano. Consideremos el problema siguiente. Sea dado un sistema numerable linealmente independiente de elementos x_n, $n = 1, 2, \ldots$, del espacio X. Se requiere obtener del sistema dado, con ayuda de combinaciones lineales finitas, un sistema ortogonal. Resulta que este problema siempre tiene solución.

Teorema 1. *Sea*

$$x_n, \quad n = 1, 2, \ldots, \tag{58.4}$$

un sistema linealmente independiente de elementos del espacio X. Existe tal sistema ortogonal de elementos y_n, $y_n \neq 0$, $n = 1, 2, \ldots$, de este espacio que cualquiera de sus elementos y_n, $n = 1, 2, \ldots$, es una combinación lineal de los primeros n elementos del sistema (58.4):

$$y_n = \alpha_{n,1}x_1 + \alpha_{n,2}x_2 + \ldots + \alpha_{n,n}x_n. \tag{58.5}$$

La construcción del sistema ortogonal $\{y_n\}$ del tipo (58.5), partiendo del sistema linealmente independiente $\{x_n\}$, se denomina, corrientemente, *proceso de ortogonalización del sistema* $\{x_n\}$.

DEMOSTRACIÓN. Pongamos $y_1 = x_1$. Por cuanto el sistema (58.4) es linealmente independiente, entonces $y_1 \neq 0$ (¿por qué?).

Supongamos que existen elementos ortogonales dos a dos $y_n \neq 0$, $n = 1, 2, \ldots, k$, $k \geqslant 1$, que satisfacen la condición (58.5). Buscaremos un elemento y_{k+1}, que sea ortogonal a todos los $y_1, \ldots, y_k$, en la forma

$$y_{k+1} = \beta_{k+1,1} y_1 + \ldots + \beta_{k+1,k} y_k - x_{k+1}. \tag{58.6}$$

De las condiciones de ortogonalidad

$$(y_1, y_{k+1}) = \ldots = (y_k, y_{k+1}) = 0 \tag{58.7}$$

obtenemos

$$(y_1, y_1)\beta_{k+1,1} = (y_1, x_{k+1}), \ldots, (y_k, y_k)\beta_{k+1,k} = (y_k, x_{k+1}). \tag{58.8}$$

De aquí se determinan unívocamente los coeficientes $\beta_{k+1,i}$, $i = 1, 2, \ldots, k$. El elemento y_{k+1}, prescrito por la representación (58.6) con los coeficientes determinados $\beta_{k+1,i}$, $i = 1, 2, \ldots, k$, satisface las condiciones (58.7).

Sustituyamos en (58.6) las expresiones para y_n, $n = 1, 2, \ldots, k$ escritas en la forma (58.5); realizada la reducción de los términos semejantes, se obtiene

$$y_{k+1} = \alpha_{k+1,1} x_1 + \ldots + \alpha_{k+1,k} x_k - x_{k+1}. \tag{58.9}$$

De aquí proviene que $y_{k+1} \neq 0$, pues de lo contrario los elementos $x_1, \ldots, x_{k+1}$ resultarían ser linealmente dependientes. □

OBSERVACIÓN. Indiquemos que si un sistema ortogonal de elementos z_n, $z_n \neq 0$, $n = 1, 2, \ldots$, del espacio X es de tal índole que todo elemento z_n es también una combinación lineal de los primeros n elementos del sistema (58.4):

$$z_n = \gamma_{n,1} x_1 + \ldots + \gamma_{n,n} x_n, \quad n = 1, 2, \ldots, \tag{58.10}$$

entonces el elemento z_n puede diferir del elemento y_n sólo en cierto factor numérico $\lambda_n \neq 0$:

$$z_n = \lambda_n y_n, \quad n = 1, 2, \ldots.$$

Demostremos esto. Designemos mediante $L(u_1, \ldots, u_n)$ la cápsula lineal del sistema de elementos $u_1, \ldots, u_n$ (véase el p. 57.2); $L(x_1, \ldots, x_n)$ será un espacio n-dimensional en el cual los elementos $x_1, \ldots, x_n$ forman la base (véase el p. 57.2). Los elementos y_i, $i = 1, 2, \ldots, n$ (z_i, $i = 1, 2, \ldots, n$, respectivamente) son linealmente independientes y están contenidos en $L = (x_1, \ldots, x_n)$; por consiguiente, los elementos y_i, $i = 1, 2, \ldots, n$ y los z_i, $i = 1, 2, \ldots, n$, forman también una base en el espacio $L(x_1, \ldots, x_n)$. De este modo, $L(x_1, \ldots, x_n) = L(y_1, \ldots, y_n) = L(z_1, \ldots, z_n)$, $n = 1, 2, \ldots$.

El elemento $y_n \in L(x_1, \ldots, x_n)$ es ortogonal al subespacio $L(y_1, \ldots, y_{n-1}) = L(x_1, \ldots, x_{n-1})$, es decir, ortogonal a todo elemento de este subespacio. Mientras tanto, el elemento $z_n \in L(x_1, \ldots, x_n)$ es ortogonal al subespacio $L(x_1, \ldots, z_{n-1}) = L(x_1, \ldots, x_{n-1})$. Así pues, los elementos y_n y z_n del espa-

cio n-dimensional $L(x_1, \ldots, x_n)$ son ortogonales a un mismo subespacio $(n-1)$-dimensional $L(x_1, \ldots, x_{n-1})$ y, por ende, son colineales: $z_n = \lambda_n y_n$, $\lambda_n \neq 0$, $n = 1, 2, \ldots$ (¿por qué?).

Diremos también que de

$$L(x_1, \ldots, x_n) = L(y_1, \ldots, y_n), \quad n = 1, 2, \ldots$$

se infiere la coincidencia de las cápsulas lineales de los sistemas *infinitos* (58.4) y (58.5).

Analicemos ahora el sistema de potencias de x:

$$1, x, x^2, \ldots, x^n, \ldots \tag{58.11}$$

Este sistema es linealmente independiente en cualquier segmento (finito o infinito). Efectivamente, si

$$\lambda_0 + \lambda_1 x + \ldots + \lambda_n x^n = 0, \tag{58.12}$$

entonces, derivando esta identidad n veces, obtendremos

$$n!\lambda_n = 0,$$

es decir, $\lambda_n = 0$.

Si ya ha sido demostrado que $\lambda_{k+1} = \ldots = \lambda_n = 0$, la identidad (59.12) toma la forma

$$\lambda_0 + \lambda_1 x + \ldots + \lambda_k x^k = 0.$$

Derivándola k veces, obtendremos $\lambda_k = 0$. Así pues, $\lambda_1 = \ldots = \lambda_n = 0$, lo que precisamente significa la independencia lineal de las funciones $1, x, \ldots, x^n$.

Ha de notarse que por cuanto las funciones del sistema (58.11), que se consideran en cierto segmento $[a, b]$, pertenecen a los espacios $C[a, b]$ (véase el ejemplo 7 en el p. 57.4), $CL_2[a, b]$ y $L_2[a, b]$ (véase el p. 57.10), entonces en dichos espacios se tienen sistemas linealmente independientes infinitos. Por consiguiente, los espacios citados son de dimensión infinita, es decir, a ciencia cierta no tienen base compuesta de un número finito de elementos.

Tomemos el sistema (58.11) en el segmento $[-1, 1]$ a título de sistema de partida (58.4) y apliquemos a (58.11) el proceso de ortogonalización (véase (58.5)) en el espacio $L_2[-1, 1]$; en este caso obtendremos una sucesión de polinomios ortogonales de grados 0, 1, 2, . . . , respectivamente. De la observación aducida más arriba se deduce que estos polinomios pueden diferir de los de Legendre (58.3), también ortogonales, sólo en un factor constante.

58.3. SISTEMAS COMPLETOS. COMPLETITUD DEL SISTEMA TRIGONOMÉTRICO Y DEL SISTEMA DE LOS POLINOMIOS DE LEGENDRE

Recordemos (véase el p. 57.6) que un sistema de elementos $\Omega = [x_\alpha]$, $\alpha \in \mathfrak{A}$, se llama *completo en el espacio seminormalizado* X, si el conjunto de todas las combinaciones lineales finitas de sus elementos es denso en el espacio X en el sentido de la seminorma prescrita en él. En otras palabras, un sistema es completo, si para todo

$x \in X$ y cualquier $\varepsilon > 0$ existen tales elementos $x_{\alpha_k} \in \Omega$ y números λ_k, $k = 1, 2, \ldots, m$, que

$$\left\| x - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\| < \varepsilon.$$

Definición 3. *Un espacio seminormalizado X se denomina encajado en el espacio seminormalizado Y, si*

1°) $X \subset Y$;

2°) *existe una constante $c > 0$ tal que para todo $x \in X$ tiene lugar la desigualdad*

$$\|x\|_Y \leqslant c\|x\|_X.$$

La constante $c > 0$ se llama *constante de encaje*. El encaje del espacio X en el espacio Y se designa mediante el símbolo

$$X \Subset Y.$$

Es fácil comprobar que si $X \Subset Y$ e $Y \Subset Z$, entonces $X \Subset Z$. Del lema 3, el p. 57.4, se desprende que para cualquier segmento tienen lugar los encajes

$$RL_p[a, b] \Subset RL_1[a, b],$$

$$RL_p[a, b] \cap S[a, b] \Subset RL_p[a, b], \quad 1 \leqslant p < +\infty.$$

Aquí, en el segundo encaje el espacio $RL_p[a, b] \cap S[a, b]$ se considera con la norma $\|\cdot\|_\infty$, es decir, con la del espacio $S[a, b]$. Si nos limitamos solamente a las funciones continuas, entonces del segundo encaje se desprende un encaje

$$C[a, b] \Subset CL_p[a, b], \quad 1 \leqslant p < +\infty. \tag{58.13}$$

De aquí, recordando que para $p = 2$ el espacio $CL_2[a, b]$ se encaja isométricamente en el espacio $L_2[a, b]$ (véase (57.52)) obtenemos un encaje más

$$C[a, b] \Subset L_2[a, b]. \tag{58.14}$$

Fijemos la atención en que en los encajes (58.13) y (58.14) los espacios que se encajan son densos en los espacios donde vienen encajados: en el caso (58.13) esto proviene de que los conjuntos de puntos de ambos espacios coinciden y en el caso (58.14), del teorema 6 del p. 57.10.

Lema 3. *Si un sistema $\Omega = \{x_\alpha\}$, $\alpha \in \mathfrak{A}$, es completo en el espacio seminormalizado X, el espacio X está encajado en el espacio seminormalizado Y y el conjunto X es denso en el espacio Y según la seminorma de éste, entonces el sistema Ω será denso en el espacio Y.*

DEMOSTRACIÓN. Tomemos un elemento arbitrario $y \in Y$ y cualquier $\varepsilon > 0$. Por ser el conjunto X denso, en el espacio X se encontrará un elemento $x \in X$ tal que

$$\|y - x\|_Y < \frac{\varepsilon}{2}.$$

Por cuanto el sistema Ω es completo en el espacio X, existe un conjunto finito de elementos $x_{\alpha_k} \in \Omega$ y números λ_k, $k = 1, 2, \ldots, m$, tales que

$$\left\| x - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\|_X < \frac{\varepsilon}{2c},$$

donde $c > 0$ es una constante del encaje $X \subset Y$. En virtud de este encaje (véase la definición 3)

$$\left\| x - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\|_Y \leqslant c \left\| x - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\|_X < \frac{\varepsilon}{2}.$$

Por eso, para el elemento y, elegido inicialmente, obtendremos

$$\left\| y - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\|_Y \leqslant \|y - x\|_Y + \left\| x - \sum_{k=1}^{m} \lambda_k x_{\alpha_k} \right\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Esto es precisamente el indicio de que el sistema Ω es denso en el espacio Y. □

EJEMPLOS. 1. Un sistema de potencias

$$1, x, x^2, \ldots, x^n, \ldots \tag{58.15}$$

es completo en los espacios $C[a, b]$, $CL_p[a, b]$, $1 \leqslant p < +\infty$, y $L_2[a, b]$ para cualquier segmento $[a, b]$. En efecto, en virtud del teorema de Weierstrass (véase el teorema 8′ en el p. 55.8), el citado sistema de potencias es completo en el espacio $C[a, b]$ que está encajado, de acuerdo con (58.14), en el espacio $L_2[a, b]$ y es denso en éste. Por ello, de conformidad con el lema 3 de este punto, el sistema de potencias (58.15) es completo en el espacio $L_2[a, b]$. Según este mismo lema, el sistema mencionado es completo también en el espacio $CL_p[a, b]$ para cualquier $p \geqslant 1$, pues, $C[a, b]$ está encajado en $CL_p[a, b]$ y es denso en él (véase (58.13)).

Subrayemos que toda base en un espacio lineal normalizado es, evidentemente, un sistema completo linealmente independiente. Lo recíproco no es cierto. Por ejemplo, aunque el sistema de potencias (58.15) forma un sistema completo linealmente independiente en el espacio de Banach $C[a, b]$, no es, sin embargo, una base en él: si una función f en el espacio $C[a, b]$ se desarrolla según el sistema de potencias (58.15), es decir, si $f(x) = \sum_{n=0}^{\infty} a_n x^n$, esto implica que la serie de potencias escrita converge uniformemente en el segmento $[a, b]$, y, por lo tanto, la función f es analítica en el intervalo (a, b). Por esta razón, a ciencia cierta, ninguna función continua en el segmento $[a, b]$ puede ser representada en la forma indicada.

2. Un sistema de los polinomios de Legendre

$$P_0(x) = 1, P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n} \tag{58.3}$$

es completo en los espacios $C[a, b]$, $CL_p[a, b]$, $1 \leqslant p < +\infty$ y $L_2[a, b]$ para cualquier segmento $[a, b]$. Esto se infiere inmediatamente de que todo polinomio $Q(x)$ es una combinación lineal de los polinomios de Legendre (véase el p. 58.1):

$$Q(x) = \sum_{k=0}^{n} \lambda_k P_k(x). \tag{58.16}$$

Por eso, si en algún espacio seminormalizado X es completo el sistema de potencias

(58.15), es decir, para todo elemento $f \in X$ y cualquier $\varepsilon > 0$ existe un polinomio $Q = Q(x)$ tal que $\|f - Q\| < \varepsilon$, entonces, en virtud de (58.16),

$$\left\| f - \sum_{k=0}^{n} \lambda_k P_k \right\| < \varepsilon.$$

Esto es precisamente el testimonio de que el sistema de los polinomios de Legendre es completo en el espacio X.

3. Designemos con $C^*[-\pi, \pi]$ un subespacio del espacio de funciones continuas $C[-\pi, \pi]$, compuesto de las funciones que en los extremos del segmento $[-\pi, \pi]$ toman valores iguales:

$$f(-\pi) = f(\pi). \tag{58.17}$$

El sistema trigonométrico

$$1, \cos x, \operatorname{sen} x, \ldots, \cos nx, \operatorname{sen} nx, \ldots, \tag{58.2}$$

es completo en los espacios $C^*[-\pi, \pi]$ y $L_2[-\pi, \pi]$. La completitud del sistema trigonométrico en el espacio $C^*[-\pi, \pi]$ ha sido demostrada anteriormente: véase el teorema 7′ en el p. 55.8.

Denotemos con $\mathring{C}[-\pi, \pi]$ un subespacio del espacio $C^*[-\pi, \pi]$, compuesto de aquellas funciones f que en los extremos del segmento $[-\pi, \pi]$ toman valores nulos: $f(-\pi) = f(\pi) = 0$. De acuerdo con el teorema 6, p. 57.10, el conjunto $\mathring{C}[-\pi, \pi]$, y, por ende, también el espacio $C^*[-\pi, \pi] \supset \mathring{C}[-\pi, \pi]$, es denso en el espacio $L_2[-\pi, \pi]$. Por esta razón, en virtud del encaje (véase (58.14))

$$C^*[-\pi, \pi] \Subset L_2[-\pi, \pi]$$

y del lema 3 de este punto, el sistema trigonométrico (58.2) es completo en el espacio $L_2[-\pi, \pi]$.

Diremos que por cuanto la condición (58.17) se conserva en el caso de convergencia uniforme y todo polinomio trigonométrico satisface a esta condición, entonces el sistema trigonométrico no es completo, a ciencia cierta, en el espacio $C[-\pi, \pi]$, puesto que éste contiene a todas luces funciones que no satisfacen la condición (58.17).

Como un corolario se desprende de los ejemplos considerados la siguiente afirmación.

Teorema 2. *El espacio de Banach $C[a, b]$ y el de Hilbert $L_2[a, b]$ son espacios separables.*

Efectivamente, la separabilidad de un espacio significa (véase la definición 31 en el p. 57.6) que en dicho espacio está presente un sistema numerable completo. En los espacios de Banach y de Hilbert en calidad de tal sistema interviene el sistema (58.15) de potencias enteras no negativas de la variable x.

58.4. SERIES DE FOURIER

Sea, al igual que hasta ahora, X un espacio prehilbertiano. Consideremos el siguiente problema. Sea dado un sistema de n vectores linealmente independientes $e_1, e_2, \ldots, e_n$ del espacio X y sea fijado un cierto vector $x \in X$. Se requiere hallar una

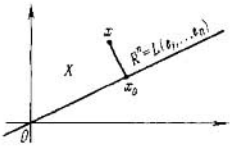


Fig. 241

combinación lineal de la forma

$$a_1e_1 + \ldots + a_ne_n, \tag{58.18}$$

la que ofrece la mejor aproximación en el espacio X del elemento x, es decir, realiza el mínimo de la expresión

$$\|x - (a_1e_1 + \ldots + a_ne_n)\|, \tag{58.19}$$

o, que es lo mismo, el mínimo de la función

$$\left\|x - \sum_{k=1}^{n} a_ke_k\right\|^2 = \left(x - \sum_{k=1}^{n} a_ke_k, x - \sum_{k=1}^{n} a_ke_k\right) \tag{58.20}$$

de las variables $a_1, \ldots, a_n$.

En el lenguaje geométrico esto significa que en un espacio n-dimensional $R^n = L(e_1, \ldots, e_n)$ tendido sobre los vectores $e_1 \in X, \ldots, e_n \in X$ se busca un elemento que sea menos alejado del elemento dado $x \in X$.

Si el espacio X es n-dimensional y, por consiguiente, los vectores $e_1, \ldots, e_n$ forman una base, siempre pueden elegirse tales coeficientes a_k, $k = 1, 2, \ldots, n$, que se cumpla la igualdad

$$x = a_1e_1 + \ldots + a_ne_n \tag{58.21}$$

y, por lo tanto, la expresión (58.19) se reduzca a cero. En cambio, si X no es de dimensión finita o es de dimensión finita, pero su dimensión es superior a n, entonces la igualdad (58.21), en el caso general, no es posible de realizarla y el problema consiste en la búsqueda de una combinación lineal (58.18) que dé el valor mínimo a la expresión (58.19).

Mostremos que el problema enunciado siempre tiene una solución x_0 que es única y, además, aclaremos algunas propiedades de esta solución (véase la fig. 241, en la que está expuesto esquemáticamente el problema en consideración). Al emplear, si es necesario, el proceso de ortogonalización (véase el p. 58.2), siempre podemos sustituir el sistema $e_1, \ldots, e_n$ por un sistema ortogonal de vectores distintos de cero. Por eso supondremos que $e_k \neq 0$, $(e_k, e_j) = 0$, $k \neq j, j, k = 1, 2, \ldots, n$. Haciendo uso de la condición de ortogonalidad, transformemos la función (58.20) del modo siguiente:

$$\left\|x - \sum_{k=1}^{n} a_ke_k\right\|^2 = \left(x - \sum_{k=1}^{n} a_ke_k, x - \sum_{j=1}^{n} a_je_j\right) =$$

$$= (x, x) + \sum_{k=1}^{n} \sum_{j=1}^{n} a_k a_j (e_k, e_j) - 2 \sum_{k=1}^{n} a_k (x, e_k) =$$

$$= \|x\|^2 + \sum_{k=1}^{n} a_k^2 \|e_k\|^2 - 2 \sum_{k=1}^{n} a_k (x, e_k) =$$

$$= \|x\|^2 + \sum_{k=1}^{n} \left(a_k \|e_k\| - \frac{(x, e_k)}{\|e_k\|} \right)^2 - \sum_{k=1}^{n} \frac{(x, e_k)^2}{\|e_k\|^2}, \quad (58.22)$$

De aquí se desprende*) que el mínimo de la expresión (58.19) se alcanza cuando

$$a_k \|e_k\| - \frac{(x, e_k)}{\|e_k\|} = 0, \quad k = 1, 2, \ldots, n,$$

es decir, cuando

$$a_k = \frac{(x, e_k)}{\|e_k\|^2}. \quad (58.23)$$

Los números a_k, definidos por la fórmula (58.23), *se denominan coeficientes de Fourier del elemento x según el sistema $e_1, \ldots, e_n$.*

Si el sistema $e_1, \ldots, e_n$ es ortonormalizado, las fórmulas (58.23) adquieren una forma más sencilla:

$$a_k = (x, e_k). \quad (58.24)$$

En el caso de un espacio n-dimensional, cuando a título de vectores $e_1, \ldots, e_n$ está elegida la base del espacio, los coeficientes de Fourier del vector x son los coeficientes de su desarrollo según la base citada, es decir, las coordenadas del elemento x respecto de esta base. Resulta fácil convencerse de esto, al multiplicar escalarmente la igualdad (58.21) por e_k, $k = 1, 2, \ldots, n$: como resultado se obtendrá (58.23).

Volveremos ahora a la expresión (58.22). Al tomar en ella, a título de $a_1, \ldots, a_n$, los coeficientes de Fourier (58.23), obtendremos

$$\|x\|^2 - \sum_{k=1}^{n} a_k^2 \|e_k\|^2 = \left\| x - \sum_{k=1}^{n} a_k e_k \right\|^2 \geqslant 0. \quad (58.25)$$

de donde

$$\sum_{k=1}^{n} a_k^2 \|e_k\|^2 \leqslant \|x\|^2. \quad (58.26)$$

Así pues, queda demostrado el siguiente teorema.

Teorema 3. *Supongamos que $e_k, e_k \neq 0, k = 1, 2, \ldots, n$, es un sistema ortogonal de vectores del espacio prehilbertiano X. La mejor aproximación en el espacio X del vector $x \in X$ por medio de las combinaciones lineales del tipo $\sum_{k=1}^{n} \alpha_k e_k$ se*

*) Es obvio que estos razonamientos son una generalización inmediata de la demostración del teorema 11 del p. 55.9.

efectúa cuando $\alpha_k, k = 1, 2, \ldots, n$, *son coeficientes de Fourier:* $\alpha_k = a_k$. *En este caso*

$$\inf_{\alpha_1, \ldots, \alpha_n} \left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\|^2 = \left\| x - \sum_{k=1}^{n} a_k e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^{n} a_k^2 \|e_k\|^2 \geqslant 0.$$

Corolario 1. *El elemento* $x_0 = \sum_{j=1}^{n} \alpha_j e_j$ *es el de mejor aproximación del elemento* $x \in X$ *en el subespacio* $L(e_1, \ldots, e_n)$ *cuando, y sólo cuando, el elemento* $x - x_0$ *sea ortogonal a* $L(e_1, \ldots, e_n)$, *es decir, cuando* $x - x_0 \perp L(e_1, \ldots, e_n)$.

Efectivamente, la condición $x - x_0 \perp L(e_1, \ldots, e_n)$ es equivalente a la otra: para todos los $k = 1, 2, \ldots, n$ se verifica la igualdad $(x - x_0, e_k) = 0$. Esta, a su vez, es equivalente a la condición $(x, e_k) = (x_0, e_k)$, o bien, por cuanto

$$(x_0, e_k) = \left(\sum_{j=1}^{n} \alpha_j e_j, e_k \right) = \alpha_k (e_k, e_k),$$

a la condición $(x, e_k) = \alpha_k (e_k, e_k)$. De este modo, las condiciones

$$x - x_0 \perp L(e_1, \ldots, e_n) \quad \text{y} \quad \alpha_k = \frac{(x, e_k)}{(e_k, e_k)}$$

son equivalentes. Mas, la segunda condición significa que los números α_k son coeficientes de Fourier del elemento x_0, es decir, x_0 es el elemento de la mejor aproximación. □

Sea dada ahora una sucesión (no el sistema finito, como antes) de elementos

$$e_n (e_n \neq 0), \quad n = 1, 2, \ldots, \tag{58.27}$$

que forman un sistema ortogonal en el espacio X. Los números $a_k, k = 1, 2, \ldots$, determinados por la fórmula (58.23) se llamarán en este caso también *coeficientes de Fourier* del elemento y según el sistema (58.27).

Definición 4. *La serie*

$$\sum_{n=1}^{\infty} a_n e_n, \tag{58.28}$$

donde $a_n, n = 1, 2, \ldots$ *son los coeficientes de Fourier* (58.23) *del elemento* x *según el sistema* (58.27), *se llama serie de Fourier del elemento* x *según dicho sistema.*

Si la serie (58.28) es la de Fourier del elemento x, se escribe

$$x \sim \sum_{n=1}^{\infty} a_n e_n.$$

Definición 5. *Sean dados un sistema ortogonal* (58.27) *y un elemento* $x \in X$. *Se denomina la mejor aproximación del elemento* x *por medio de las combinaciones lineales del tipo* $\sum_{k=1}^{n} \alpha_k e_n$ (n *está fijo*) *al número* $E_n(x)$ *determinado por la igualdad*

$$E_n(x) = \inf_{\alpha_1, \ldots, \alpha_n} \left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\|, \quad n = 1, 2, \ldots,$$

donde la cota inferior se toma según toda clase de coeficientes $\alpha_1, \ldots, \alpha_n$, *o, que es lo mismo, según toda clase de combinaciones lineales del tipo* $\sum_{k=1}^{n} \alpha_k e_k$.

Por cuanto toda combinación lineal de elementos $e_1, \ldots, e_n$ puede considerarse también como una combinación lineal de elementos $e_1, \ldots, e_n, e_{n+1}$, entonces, evidentemente,

$$E_{n+1}(x) \leqslant E_n(x). \tag{58.29}$$

Del teorema 3 se infiere que la cota inferior de que se trata se alcanza, si a título de coeficientes α_k se toman los coeficientes de Fourier y que

$$E_n(x) = \inf_{\alpha_1, \ldots, \alpha_n} \left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\| =$$

$$= \left\| x - \sum_{k=1}^{n} a_k e_k \right\| = \sqrt{\|x\|^2 - \sum_{k=1}^{n} a_k^2 \|e_k\|^2},$$

$$a_k = \frac{(x, e_k)}{(e_k, e_k)}, \quad k = 1, 2, \ldots. \tag{58.30}$$

El resultado obtenido lo enunciemos en forma del corolacio del teorema 3.

Corolario 2. *Las sumas parciales*

$$s_n = \sum_{k=1}^{n} a_k e_k$$

de la serie de Fourier del elemento $x \in X$ *realizan la mejor aproximación en el espacio* X *del elemento* $x \in X$ *por medio de las combinaciones lineales del tipo* $\alpha_1 e_1 + \ldots + \alpha_n e_n$.

Demos a conocer algunos otros corolarios del teorema 3.

Corolario 3. *Si* s_n *es la suma parcial de la serie de Fourier del elementos* $x \in X$, *la sucesión numérica* $\|x - s_n\|$ *decrece*:

$$\|x - s_{n+1}\| \leqslant \|x - s_n\|, \quad n = 1, 2, \ldots \tag{58.31}$$

Efectivamente, de acuerdo con (58.30),

$$\|x - s_n\| = E_n(x), \quad n = 1, 2, \ldots$$

Por ello la desigualdad (58.31) es, de hecho, la desigualdad (58.29) escrita en otras designaciones.

Corolario 4. *Para los coeficientes de Fourier* a_n, $n = 1, 2, \ldots$, *de todo elemento* $x \in X$ *se verifica la desigualdad*

$$\sum_{n=1}^{\infty} a_n^2 \|e_n\|^2 \leqslant \|x\|^2 \tag{58.32}$$

llamada desigualdad de Bessel.

La desigualdad (58.32) proviene directamente de la (58.26) cuando $n \to \infty$ (compárese con la desigualdad (55.49) en el p. 55.9).

Corolario 5. *Si existe una constante $c > 0$ tal que $\|e_n\| \geqslant c$ cuando $n = 1, 2, \ldots$, en particular, si el sistema* (58.27) *es ortonormalizado (en este caso se puede tomar $c = 1$), entonces los coeficientes de Fourier de cualquier elemento $x \in X$ tienden a cero cuando $n \to \infty$*

$$\lim_{n \to \infty} a_n = 0. \tag{58.33}$$

Esto se deduce de la convergencia de la serie

$$\sum_{n=1}^{\infty} a_n^2 \leqslant \frac{1}{c^2} \sum_{n=1}^{\infty} a_n^2 \|e_n\|^2 \leqslant \frac{\|x\|^2}{c^2},$$

pues el término general de una serie convergente tiende a cero.

Surge naturalmente una cuestión: ¿bajo qué condiciones converge la serie de Fourier del elemento x?

Teorema 4. *Si el espacio X es de Hilbert (es decir, es completo), la serie de Fourier* (58.28) *de cualquier elemento $x \in X$ según todo sistema ortogonal* (58.27) *converge en el espacio X. Si x_0 es la suma de la serie:*

$$x_0 = \sum_{n=1}^{\infty} a_n e_n, \tag{58.34}$$

el elemento $x - x_0$ es ortogonal a todos los elementos del sistema (58.27).

DEMOSTRACIÓN. Sean $s_n = \sum_{k=1}^{n} a_k e_k$, $n = 1, 2, \ldots$, las sumas parciales de la serie de Fourier (58.28) del elemento x según el sistema (58.27); entonces

$$\|s_{n+p} - s_n\|^2 = \left\| \sum_{k=n+1}^{n+p} a_k e_k \right\|^2 =$$

$$= \left(\sum_{k=n+1}^{n+p} a_k e_k, \sum_{k=n+1}^{n+p} a_k e_k \right) = \sum_{k=n+1}^{n+p} a_k^2 \|e_k\|^2,$$

$$n = 1, 2, \ldots, p = 1, 2, \ldots. \tag{58.35}$$

En virtud de la desigualdad de Bessel (58.32) la serie

$$\sum_{n=1}^{\infty} a_n^2 \|e_n\|^2$$

converge y, por lo tanto, en virtud del criterio de Cauchy para la convergencia de una serie numérica, existe, para cada número $\varepsilon > 0$, tal número n_ε que, siendo $n \geqslant n_\varepsilon$ y $p > 0$, se verifica la desigualdad

$$\sum_{k=n+1}^{n+p} a_k^2 \|e_k\|^2 < \varepsilon^2,$$

por lo cual, de acuerdo con la desigualdad (58.35) para $n \geqslant n_\varepsilon$ y $p > 0$, tenemos

$$\|s_{n+p} - s_n\| < \varepsilon,$$

es decir, la sucesión $[s_n]$ es fundamental en el espacio X y, por cuanto este último es completo, converge.

En las condiciones del teorema la sucesión s_n converge, en el caso general, no hacia el elemento x. Supongamos que su límite es el elemento x_0, es decir, $x_0 = \sum_{n=1}^{\infty} a_n e_n$, entonces, haciendo uso de la continuidad del producto escalar (véase el p. 57.9) y la fórmula (58.23), obtendremos

$$(x - x_0, e_k) = (x, e_k) - (x_0, e_k) =$$

$$= (x, e_k) - \left(\sum_{n=1}^{\infty} a_n e_n, e_k\right) = (x, e_k) - \sum_{n=1}^{\infty} a_n (e_n, e_k) =$$

$$= (x, e_k) - a_k \|e_k\|^2 = 0, \quad k = 1, 2, \ldots. \square$$

En lo que se refiere a la condición de convergencia de la serie de Fourier *de cierto elemento suelto* hacia este mismo elemento, puede formularse en la forma siguiente.

Teorema 5. *La serie de Fourier* (58.28) *del elemento x de un espacio prehilbertiano converge hacia este elemento cuando, y sólo cuando, para dicha serie se cumple la igualdad*

$$\|x\|^2 = \sum_{n=1}^{\infty} a_n^2 \|e_n\|^2, \tag{58.36}$$

donde a_n son coeficientes de Fourier del elemento x según el sistema (58.27).

La igualdad (58.36) lleva el nombre de *Parseval*.

En el caso en que el sistema (58.27) sea ortonormalizado, la igualdad de Parseval toma una forma más sencilla:

$$\|x\|^2 = \sum_{n=1}^{\infty} a_n^2, \; a_n = (x, e_n), \quad n = 1, 2, \ldots,$$

y representa en sí una generalización del teorema de Pitágoras a los espacios de dimensión infinita.

DEMOSTRACIÓN DEL TEOREMA 5. Tuvimos (véase (58.25))

$$\left\|x - \sum_{k=1}^{n} a_k e_k\right\|^2 = \|x\|^2 - \sum_{k=1}^{n} a_k^2 \|e_k\|^2.$$

Pasando aquí al límite para $n \to \infty$, obtendremos la equivalencia de la condición

$$\lim_{n \to \infty} \left\|x - \sum_{k=1}^{n} a_k e_k\right\| = 0 \tag{58.37}$$

y de la condición

$$\lim_{n \to \infty} \left(\|x\|^2 - \sum_{k=1}^{n} a_k^2 \|e_k\|^2\right) = 0,$$

es decir, de la condición

$$\|x\|^2 = \lim_{n \to \infty} \sum_{k=1}^{n} a_k^2 \|e_k\|^2. \quad \square \tag{58.38}$$

Recordemos ahora el concepto de sistema completo (véase el p. 57.6), aplicado sólo al caso de sistemas numerables. El sistema de elementos $e_n \in X$, $n = 1, 2, \ldots$, se llama completo, si el conjunto de combinaciónes lineales finitas de elementos de este sistema es denso en el espacio X. Esto significa que para todo elemento $x \in X$ y todo número $\varepsilon > 0$ existe un número $n = n(\varepsilon, x)$ y los números $\lambda_1, \ldots, \lambda_n$ tales que se verifica la desigualdad

$$\|x - (\lambda_1 e_1 + \ldots + \lambda_n e_n)\| < \varepsilon. \tag{58.39}$$

La completitud del sistema ortonormalizado constituye una condición que asegura la convergencia de la serie de Fourier de *cualquier elemento* del espacio hacia este mismo elemento. Enunciemos esta condición en forma de un teorema.

Teorema 6. *Una serie de Fourier de cualquier elemento de un espacio prehilbertiano según el sistema ortogonal* (58.27) *converge hacia este mismo elemento cuando, y sólo cuando, el sistema* (58.27) *es completo.*

Corolario. *Para que el sistema ortogonal* (58.27) *de un espacio prehilbertiano X sea completo en el espacio X, es necesario y suficiente que para cualquier elemento $x \in X$ se verifique la igualdad de Parseval* (58.36).

DEMOSTRACIÓN DEL TEOREMA 6. Sea X un espacio prehilbertiano y supongamos que el sistema (58.27) es el sistema ortogonal de dicho espacio. Si para cualquier $x \in X$ su serie de Fourier según el sistema (58.27) converge hacia x, es decir,

$$x = \sum_{n=1}^{\infty} a_n e_n, \quad \text{donde} \quad a_n = \frac{(x, e_n)}{\|e_n\|^2}, \quad n = 1, 2, \ldots, \tag{58.40}$$

entonces

$$\lim_{n \to \infty} \left\| x - \sum_{k=1}^{n} a_k e_k \right\| = 0. \tag{58.41}$$

Por consiguiente, para cada número $\varepsilon > 0$ existe tal suma parcial $s_n = \sum_{k=1}^{n} a_k e_k$ de la serie de Fourier (58.28), que

$$\|x - s_n\| < \varepsilon, \tag{58.42}$$

es decir, se cumple la condición (58.39)

Viceversa, si la condición (58.39) se cumple para algunos coeficientes $\lambda_1, \ldots, \lambda_n$, se cumple también a ciencia cierta, de acuerdo con el teorema 3, en el caso en que se toma $\lambda_1 = a_1, \ldots, \lambda_n = a_n$, es decir, en este caso para $\varepsilon > 0$ dado se cumple la condición (58.42) con cierto n, y, por endè; con cualesquiera $m > n$ (véase (58.31)), lo que es equivalente al cumplimiento de la condición (58.41). $\square$

El corolario proviene directamente de los teoremas 5 y 6.

Aclaremos ahora la cuestión sobre la unicidad de un elemento que en calidad de su serie de Fourier tiene la serie dada $\sum_{n=1}^{\infty} a_n e_n$.

Teorema 7. *Si el sistema ortogonal* (58.27) *de un espacio prehilbertiano X es completo, el elemento $x \in X$, cuyos coeficientes de Fourier según el sistema* (58.27) *son todos iguales a cero, es por sí mismo nulo.*

Corolario. *Si dos elementos del espacio X según el sistema ortogonal completo* (58.27) *tienen todos los coeficientes de Fourier iguales entonces son iguales los propios elementos.*

DEMOSTRACIÓN DEL TEOREMA 7. Si el sistema (58.27) es completo, entonces, de acuerdo con el teorema 6, cualquier elemento $x \in X$ interviene como suma de su serie de Fourier: $x = \sum_{n=1}^{\infty} a_n e_n$. Por ello, si $a_n = 0$, $n = 1, 2, \ldots$, entonces también $x = 0$.

DEMOSTRACIÓN DEL COROLARIO. Si $x_1 \in X$, $x_2 \in X$ y si, además, sus coeficientes de Fourier son iguales entre sí:

$$\frac{(x_1, e_n)}{\|e_n\|^2} = \frac{(x_2, e_n)}{\|e_n\|^2}, \quad n = 1, 2, \ldots,$$

entonces para el elemento $x = x_1 - x_2$ todos los coeficientes de Fourier son nulos:

$$\frac{(x, e_n)}{\|e_n\|^2} = \frac{(x_1 - x_2, e_n)}{\|e_n\|^2} = \frac{(x_1, e_n)}{\|e_n\|^2} - \frac{(x_2, e_n)}{\|e_n\|^2} = 0, \quad n = 1, 2, \ldots,$$

y, por lo tanto, de acuerdo con el teorema, $x = 0$, es decir, $x_1 = x_2$. □

OBSERVACIÓN. Conviene notar que si en un espacio prehilbertiano X está dado cierto sistema ortogonal $\{e_n\}$, $e_n \neq 0$, y para cierto $x \in X$ existe una representación suya en la forma

$$x = \sum_{n=1}^{\infty} x_n e_n,$$

será única y los coeficientes x_n, $n = 1, 2, \ldots$, son los de Fourier. En efecto, si la representación citada existe, para todo $m = 1, 2, \ldots$, obtenemos, en virtud de la ortogonalidad del sistema $\{e_n\}$:

$$(x, e_m) \doteq \left(\sum_{n=1}^{\infty} x_n e_n, e_m \right) = \sum_{n=1}^{\infty} x_n (e_n, e_m) = x_m (e_m, e_m),$$

de donde

$$x_m = \frac{(x, e_m)}{(e_m, e_m)},$$

es decir, los coeficientes x_m, $m = 1, 2, \ldots$ se determinan unívocamente y coinciden con los coeficientes de Fourier.

Así pues, si en un espacio prehilbertiano se tiene un sistema ortogonal completo, todo elemento de este espacio se desarrolla en serie según dicho sistema (teorema 6)

y, además, de conformidad con la observación citada, de modo único. En otras palabras (véase la definición 33 en el p. 57.6), *todo sistema ortogonal completo* $\{e_n\}$, $e_n \neq 0$, $n = 1, 2, \ldots$, *en particular, todo sistema completo ortonormalizado de un espacio prehilbertiano constituye la base de éste.*

Por ejemplo, con arreglo a los resultados del p. 58.3, los polinomios de Legendre (58.3) forman la base en el espacio de Hilbert $L_2[-1, 1]$, mientras que el sistema trigonométrico (58.2), la base en el espacio de Hilbert $L_2[-\pi, \pi]$.

Demos a conocer ahora otro enfoque al concepto de completitud del sistema ortogonal en un espacio completo.

Definición 6. *El sistema ortogonal* (58.27) *se denomina cerrado, si en el espacio* X *no existe un elemento distinto de cero y ortogonal a cualquiera de los elementos del sistema* (58.27).

Teorema 8. *Si el espacio* X *es completo, el sistema ortogonal* (58.27) *será completo cuando, y sólo cuando, sea cerrado.*

DEMOSTRACIÓN. Si el sistema (58.27) es completo, $x \in X$ y x es ortogonal a todos los elementos del sistema (58.27), entonces todos sus coeficientes de Fourier según el sistema (58.27) son nulos (véase (58.23)), por consiguiente (teorema 7), $x = 0$.

Viceversa, sea el sistema (58.27) cerrado, $x \in X$ y $x \sim \sum_{n=1}^{\infty} a_n e_n$. De acuerdo con el teorema 4, la serie de Fourier del elemento x converge, y si $x_0 = \sum_{n=1}^{\infty} a_n e_n$, entonces $x - x_0 \perp e_n$, $n = 1, 2, \ldots$. Por ello, debido a que el sistema (58.27) es cerrado, $x - x_0 = 0$, es decir, $x = x_0$ y $x = \sum_{n=1}^{\infty} a_n e_n$. Por cuanto x es un elemento arbitrario del espacio X, de aquí proviene, en virtud del teorema 6, la completitud del espacio (58.27). □

Problema 39. Aclárese si son equivalentes o no el concepto de sistema ortogonal completo y el de sistema ortogonal cerrado en cualquier espacio prehilbertiano.

58.5. EXISTENCIA DE LA BASE EN LOS ESPACIOS SEPARABLES DE HILBERT. ISOMORFISMO DE LOS ESPACIOS SEPARABLES DE HILBERT

Teorema 9. *En todo espacio lineal separable* X *provisto de un producto escalar existe una base ortonormalizada* e_n, $n = 1, 2, \ldots$.

DEMOSTRACIÓN. Cuando el espacio X es n-dimensional, el teorema es evidente (véase el p. 18.4 y p. 57.2), razón por la cual se considerará sólo el caso en que X es de dimensión infinita.

Por cuanto X es un espacio separable, existen en él las sucesiones de elementos

$$\varphi_n, \quad n = 1, 2, \ldots,$$

que forman un sistema completo. Desechando sucesivamente aquellos de los elementos que son combinaciones lineales de los restantes, obtendremos una sucesión de elementos

$$\psi_n, \quad n = 1, 2, \ldots,$$

los cuales tienen la misma cápsula lineal que el sistema de partida $\{\varphi_n\}$ y que son, además, linealmente independientes (¿por qué?). Al aplicar al sistema obtenido el proceso de ortogonalización (véase el p. 58.2) y el de normalización (véase el p. 58.1), obtendremos un sistema ortonormalizado

$$e_k, \quad \|e_k\| = 1, \quad k = 1, 2, \ldots,$$

con la misma cápsula lineal que tiene el sistema $\{\psi_n\}$, y, por consiguiente, también la del sistema $\{\varphi_n\}$. Por cuanto, en virtud de la completitud del sistema $\{\varphi_n\}$, esta cápsula lineal es densa en X, el sistema $\{e_n\}$ es completo. Entre tanto, en el punto antecedente (véase la observación que sigue el teorema 7) se ha probado que todo sistema ortonormalizado completo de elementos de un espacio prohilbertiano constituye la base de dicho espacio. □

Teorema 10. *Todos los espacios separables de Hilbert de dimensión infinita son isomorfos entre sí*[*].

Demostremos previamente dos lemas. El primero de ellos generaliza la igualdad de Parseval (58.36).

Lema 4. *Supongamos que X es un espacio prehilbertiano y $e_n (e_n \neq 0)$, $n = 1, 2, \ldots$, un sistema ortogonal completo en X, $x \in X$, $y \in X$, y sea*

$$x \sim \sum_{n=1}^{\infty} a_n e_n, \quad y \sim \sum_{n=1}^{\infty} b_n e_n;$$

entonces

$$(x, y) = \sum_{n=1}^{\infty} a_n b_n \|e_n\|^2, \tag{58.43}$$

en particular, si se supone adicionalmente que $\|e_n\| = 1$, $n = 1, 2, \ldots$, se tiene

$$(x, y) = \sum_{n=1}^{\infty} a_n b_n.$$

La fórmula (58.43) generaliza, obviamente, la fórmula para el producto escalar en un espacio de dimensión finita (véase el p. 18.4).

DEMOSTRACIÓN. Según la definición de los coeficientes de Fourier

$$a_k = \frac{(x, e_k)}{\|e_k\|^2}, \quad b_k = \frac{(y, e_k)}{\|e_k\|^2},$$

por lo que tenemos

$$\left(x - \sum_{k=1}^{n} a_k e_k, y - \sum_{k=1}^{n} b_k e_k\right) = (x, y) - \sum_{k=1}^{n} b_k (x, e_k) -$$

$$- \sum_{k=1}^{n} a_k (y, e_k) + \sum_{k=1}^{b} a_k b_k (e_k, e_k) = (x, y) - \sum_{k=1}^{n} a_k b_k \|e_k\|^2. \tag{58.44}$$

---

[*] Véase en el p. 57.2 la definición de espacios de dimensión infinita y en el p. 57.9 (definición 36), la de isomorfismo de los espacios.

Por ser completo el sistema e_n, $n = 1, 2, \ldots$, tenemos

$$\lim_{n \to \infty} \left(x - \sum_{k=1}^{n} a_k e_k \right) = 0, \quad \lim_{n \to \infty} \left(y - \sum_{k=1}^{n} b_k e_k \right) = 0,$$

por lo cual, debido a la continuidad del producto escalar para $n \to \infty$, el primer miembro de la igualdad (58.44) tiende a cero, por consiguiente, lo mismo subsiste también para el segundo, es decir,

$$\lim_{n \to \infty} \sum_{k=1}^{n} a_k b_k \|e_k\|^2 = (x, y).$$

Esto es equivalente a la igualdad (58.43). □

Lema 5. *Supongamos que X es un espacio de Hilbert y e_k, $k = 1, 2, \ldots$, una base ortonormalizada en X; a_k, $k = 1, 2, \ldots$ una sucesión de los números tales que la serie $\sum_{k=1}^{\infty} a_k^2$ converge. En este caso la serie $\sum_{k=1}^{\infty} a_k e_k$ converge en el espacio X, y si $x = \sum_{k=1}^{\infty} a_k e_k$, entonces a_k, $k = 1, 2, \ldots$ son los coeficientes de Fourier del elemento x.*

DEMOSTRACIÓN. Si $s_n = \sum_{k=1}^{n} a_k e_k$, se tiene

$$\|s_{p+p} - s_n\|^2 = \left(\sum_{k=n+1}^{n+p} a_k e_k, \sum_{k=n+1}^{n+p} a_k e_k \right) =$$

$$= \sum_{k=n+1}^{n+p} a_k^2, \quad n = 1, 2, \ldots, \quad p = 1, 2, \ldots,$$

y, como la serie $\sum_{n=1}^{\infty} a_n^2$ es convergente, ella satisface el criterio de Cauchy para las series convergentes. De aquí proviene que la sucesión $\{s_n\}$ es fundamental en el espacio X y, por lo tanto, converge.

Sea

$$x = \lim_{n \to \infty} s_n, \quad \text{es decir} \quad x = \sum_{n=1}^{\infty} a_n e_n;$$

entonces, en virtud de que el desarrollo del elemento de un espacio según la base (véase la observación del teorema 7) es único,

$$(x, e_n) = a_n, \quad n = 1, 2, \ldots,$$

es decir, a_n son los coeficientes de Fourier del elemento x. □

DEMOSTRACIÓN DEL TEOREMA 10. Sean X e Y dos espacios separables de Hilbert de dimensión infinita. De acuerdo con el teorema 9, existen en ellos las bases ortonormalizadas e_n, $n = 1, 2, \ldots$, y f_n, $n = 1, 2, \ldots$, respectivamente.

Sea $x \in X$ y $x = \sum_{n=1}^{\infty} a_n e_n$, entonces a_n son los coeficientes de Fourier del elemento x, y, por consiguiente, según la desigualdad de Parseval, la serie $\sum_{n=1}^{\infty} a_n^2$ converge. Pongamos $y = \sum_{n=1}^{\infty} a_n f_n$. Teniendo presente el lema 5, esto tiene sentido.

La aplicación del espacio X en el espacio Y, la cual a todo elemento $x \in X$ le pone en correspondencia el elemento citado $y \in Y$, es precisamente una aplicación que realiza el isomorfismo de dichos espacios. Efectivamente, en las condiciones de esta correspondencia, a los distintos elementos del espacio X les corresponden diferentes elementos del espacio Y, puesto que el desarrollo de los elementos según la base es único. Luego, todo elemento del espacio Y está puesto en correspondencia a cierto elemento del espacio X (es decir, la citada aplicación es una aplicación sobre el espacio Y); en efecto, si $y \in Y$, entonces, al desarrollarlo en Y según la base, obtendremos

$$y = \sum_{n=1}^{\infty} b_n f_n.$$

Sea $x = \sum_{n=1}^{\infty} b_n e_n$ (tal elemento existe, véase el lema 5). Es evidente que precisamente al elemento x corresponde, en la correspondencia establecida, el elemento y. Mostremos por fin que en dicha correspondencia se conserva el producto escalar. Esto proviene inmediatamente del lema 4. Efectivamente, si

$$x = \sum_{n=1}^{\infty} a_n e_n, \quad x' = \sum_{n=1}^{\infty} b_n e_n, \quad y = \sum_{n=1}^{\infty} a_n f_n, \quad y' = \sum_{n=1}^{\infty} b_n f_n,$$

entonces, en virtud del lema mencionado,

$$(x, x') = \sum_{n=1}^{\infty} a_n b_n = (y, y'). \square$$

En calidad de modelo de un espacio separable de Hilbert de dimensión infinita puede intervenir un espacio, cuyos elementos constituyen las sucesiones de números reales

$$x = (x_1, x_2, \ldots, x_n, \ldots),$$

para las cuales la serie $\sum_{k=1}^{\infty} x_k^2$ converge, es decir, el espacio l_2 (véase el ejemplo 5 en el p. 57.4). El producto escalar en este espacio se introduce según la siguiente regla:

si $x = (x_1, \ldots, x_n, \ldots)$ *e* $y = (y_1, \ldots, y_n, \ldots)$,

entonces $(x, y) = \sum_{k=1}^{\infty} x_k y_k$.

Esta definición tiene sentido, pues de la convergencia de las series $\sum_{k=1}^{\infty} x_k^2$ y $\sum_{k=1}^{\infty} y_k^2$ se desprende que la serie $\sum_{k=1}^{\infty} x_k y_k$ es también convergente. Esto se deduce, por ejemplo, de la desigualdad de Hölder para las series cuando $p = 2$ (en este caso dicha desigualdad lleva, a menudo, el nombre de Cauchy-Schwarz), pero puede obtenerse también de la desigualdad elemental

$$x_k y_k \leqslant \frac{x_k^2 + y_k^2}{2}.$$

La norma en el espacio l_2 se determina de confirmidad con la regla general mediante la fórmula

$$\|x\| = \sqrt{\sum_{k=1}^{\infty} x_k^2}.$$

Teorema 11. *El espacio l_2 es un espacio separable de Hilbert.*

DEMOSTRACIÓN. El espacio l_2 es separable, pues las sucesiones e_k, $k = 1, 2, \ldots$, en las que todos los lugares están ocupados por ceros, a excepción del k-ésimo, donde figura la unidad, forman una base ortonormalizada y, por consiguiente, sus combinaciones lineales finitas con los coeficientes racionales forman un conjunto numerable denso en el espacio l_2 (¿por que?).

La completitud del espacio l_2 se demuestra de una manera un tanto más compleja. Sea

$$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}, \ldots), \quad k = 1, 2, \ldots, \tag{58.45}$$

una sucesión fundamental del espacio l_2. Entonces, de la desigualdad

$$\|x^{(k+p)} - x^{(k)}\| = \sqrt{\sum_{n=1}^{\infty} (x_n^{(k+p)} - x_n^{(k)})^2} \geqslant$$

$$\geqslant |x_n^{(k+p)} - x_n^{(k)}|, \quad k = 1, 2, \ldots; \; p = 0, 1, 2, \ldots, \quad n = 1, 2, \ldots,$$

y de la sucesión fundamental (58.45) se deduce que para cualquier n fijo la sucesión numérica $x_n^{(k)}$, $k = 1, 2, \ldots$, satisface el criterio de Cauchy (véase el p. 4.7) y, por consiguiente, converge. Sea $\lim_{k \to \infty} x_n^{(k)} = x_n$. Por ser la sucesión (58.45) fundamental, para todo $\varepsilon > 0$ existe tal número k_ε que se verifica la desigualdad

$$\|x^{(k+p)} - x^{(k)}\| < \varepsilon,$$

cualesquiera que sean el número $k \geqslant k_\varepsilon$ y p natural, es decir,

$$\sqrt{\sum_{n+1}^{\infty} (x_n^{(k+p)} - x_n^{(k)})^2} < \varepsilon.$$

De aquí para todo número natural fijo m tenemos con mayor razón

$$\sum_{n=1}^{m} (x_n^{(k+p)} - x_n^{(k)})^2 < \varepsilon^2.$$

Pasando aquí al límite para $p \to \infty$, obtendremos

$$\sum_{n=1}^{m} (x_n - x_n^{(k)})^2 \leqslant \varepsilon^2,$$

y, como esto es cierto para cualquier $m = 1, 2, \ldots$, se tiene

$$\sum_{n=1}^{\infty} (x_n - x_n^{(k)})^2 \leqslant \varepsilon^2, \quad k \geqslant k_\varepsilon. \tag{58.46}$$

De este modo, el punto $y^{(k)} = (x_1 - x_1^{(k)}, \ldots, x_n - x_n^{(k)}, \ldots)$, $k \geqslant k_\varepsilon$, pertenece al espacio l_2, mas, en este caso, también le pertenece a l_2 el punto $x = (x_1, \ldots, x_n, \ldots) = x^{(k)} + y^{(k)}$, mientras que la condición (58.46) significa que

$$\lim_{k \to \infty} x^{(k)} = x.$$

Hemos demostrado pues que la sucesión (58.46) converge. Por consiguiente, l_2 es un espacio completo. □

En virtud del teorema 10, el espacio l_2 es isomorfo a todo espacio separable de Hilbert.

En el p. 58.3 se ha mostrado que el espacio $L_2[a, b]$ es separable (véase allí el teorema 2) para cualquier segmento $[a, b]$, por consiguiente, es también isomorfo al espacio l_2. Se puede probar que el espacio $L_2(G)$, donde G es un conjunto medible de medida positiva del espacio n-dimensional, también es separable y, por ende, isomorfo a l_2. De este modo, todos los espacios de Hilbert de funciones integrables en cuadrado son isomorfos entre sí, independientemente del número de variable de las cuales dependen las funciones citadas.

58.6. DESARROLLO DE LAS FUNCIONES DE CUADRADO INTEGRABLE EN SERIE DE FOURIER

En el § 55 se estudiaban las series clásicas de Fourier, es decir, las series de Fourier según el sistema trigonométrico de funciones, para las funciones absolutamente integrables. En el presente punto se obtendrá toda una serie de corolarios, provenientes de la teoría general de las series de Fourier en los espacios de Hilbert y de la propiedad de completitud de un sistema de funciones trigonométricas en el espacio $L_2[-\pi, \pi]$, para las series trigonométricas de Fourier de una clase más reducida de funciones, en comparación con las funciones absolutamente integrables, a saber, para las funciones de cuadrado integrable en el segmento $[-\pi, \pi]$, es decir, para las funciones del espacio $RL_2[-\pi, \pi]$ (véase el ejemplo 3 en el p. 57.8).

Observemos, ante todo, que si en el espacio de Hilbert $L_2[-\pi, \pi]$ se toma, a título de sistema ortogonal, el sistema trigonométrico

$$1, \cos x, \operatorname{sen} x, \ldots, \cos nx, \operatorname{sen} nx, \ldots, \tag{58.2}$$

los coeficientes de Fourier del elemento $f \in L_2[-\pi, \pi]$ según este sistema se determinarán, de acuerdo con (58.23), según las fórmulas

$$a_0 = \frac{1}{2\pi}(f, 1), \quad a_n = \frac{1}{\pi}(f, \cos nx), \quad b_n = \frac{1}{\pi}(f, \operatorname{sen} nx), \tag{58.47}$$

$$n = 1, 2, \ldots$$

pues $\|1\|_{L_2} = \sqrt{2\pi}$, $\|\cos nx\|_{L_2} = \|\operatorname{sen} nx\|_{L_2} = \sqrt{\pi}$ (véase el p. 58.1).

Si f es una función continua en el segmento $[-\pi, \pi]$, entonces $f \in CL_2[-\pi, \pi] \subset L_2[-\pi, \pi]$. Comparando las fórmulas (58.47) para los coeficientes de Fourier de la función f con las (55.6) (el producto escalar se define, como siempre, mediante las fórmulas (57.30)), vemos que todas ellas coinciden, salvo la fórmula para el coeficiente a_0, la cual en (58.47) difiere de la en (55.6) en el factor 1/2. Pagando el tributo a la tradición, nos atendremos, en lo que sigue, a la fórmula (55.6) para a_0, es decir, se considerará que

$$a_0 = \frac{1}{\pi}(f, 1) \tag{58.48}$$

y la serie trigonométrica de Fourier se escribirá en la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx.$$

Aplicando el teorema 6 al sistema trigonométrico (58.2) y teniendo presente que dicho sistema es completo en el espacio $L_2[-\pi, \pi]$ (véase el ejemplo 3 en el p. 58.3), obtendremos el siguiente teorema.

Teorema 12. *Cualquier elemento $f \in L_2[-\pi, \pi]$ se desarrolla en este espacio en una serie de Fourier según el sistema trigonométrico*

$$f \simeq \frac{a_0}{2} + \sum_{n=0}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx, \tag{58.49}$$

siendo en este caso válida la igualdad de Parseval

$$\frac{1}{\pi}\|f\|^2 = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2.$$

Corolario 1. *Toda función $f(x)$ de cuadrado integrable en el sengmento $[-\pi, \pi]$ 1) es el límite, en el sentido de la media cuadrática (véase el p. 57.5), de sus sumas parciales de Fourier $S_n(x)$ según el sistema trigonométrico de funciones para $n \to \infty$, es decir,*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = 0; \tag{58.50}$$

2) y para ella es lícita la igualdad de Parseval

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x)dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2. \tag{58.51}$$

Corolario 2. *Si la función f de cuadrado integrable en el segmento $[-\pi, \pi]$ y todos sus coeficientes de Fourier según el sistema trigonométrico (58.2) son iguales a cero, la función es equivalente a cero.*

Aquí los coeficientes de Fourier para $n = 1, 2, \ldots$ se determinan siempre a partir de las fórmulas (58.47), y el coeficiente a_0, según la fórmula (58.48).

Por cuanto el propio teorema 12 se infiere del teorema 6, la demostración se necesita sólo para los corolarios.

Así pues, sea $f(x)$ una función de cuadrado integrable en el segmento $[-\pi, \pi]$, es decir, $f(x) \in RL_2[-\pi, \pi]$ (véase el ejemplo 8 en el p. 57.4 y el ejemplo 3 en el p. 57.8). Hemos de notar, ante todo, que cualquier función $g(x)$, equivalente a $f(x)$ (véase la definición 38 en el p. 57.10), tiene los mismos coeficientes de Fourier y, por consiguiente, la misma serie de Fourier. Esto proviene de que el producto semiescalar en el espacio $RL_2[-\pi, \pi]$ no varía, si sus factores se sustituyen por los equivalentes (véase la fórmula (57.41)), y por esta razón, si $f \sim g$, se tiene

$$a_0 = \frac{1}{\pi}(f, 1)_{RL_2} = \frac{1}{\pi}(g, 1)_{RL_2},$$

$$a_n = \frac{1}{\pi}(f, \cos nx)_{RL_2} = \frac{1}{\pi}(g, \cos nx)_{RL_2},$$

$$b_n = \frac{1}{\pi}(f, \operatorname{sen} nx)_{RL_2} = \frac{1}{\pi}(g, \operatorname{sen} nx)_{RL_2},$$

$$n = 1, 2, \ldots .^{*)}$$

Por consiguiente, si designamos con F la clase de funciones equivalentes en la que está contenida la función f, entonces, debido a la definición (57.41) de producto escalar de las clases de funciones equivalentes, es decir, de producto escalar en el espacio $\widetilde{RL}_2[-\pi, \pi]$ (véase el p. 57.10), tendremos

$$a_0 = \frac{1}{\pi}(F, 1)_{\widetilde{RL}_2}, \quad a_n = \frac{1}{\pi}(F, \cos nx)_{\widetilde{RL}_2}, \quad b_n = \frac{1}{\pi}(F, \operatorname{sen} nx)_{\widetilde{RL}_2}, \quad n = 1, 2, \ldots,$$

es decir, la serie de Fourier del elemento $F \in \widetilde{RL}_2[-\pi, \pi] \subset L_2[-\pi, \pi]$ coincide con la serie de Fourier de toda función $f \in F$. De acuerdo con el teorema 12, en el espacio $L_2[-\pi, \pi]$ tiene lugar el desarrollo

$$F = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx, \tag{58.52}$$

y se verifica la igualdad de Parseval

$$\frac{1}{\pi}\|F\|^2_{L_2} = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2. \tag{58.53}$$

Si $S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \operatorname{sen} kx$ es una suma parcial de la serie de Fourier (58.52), entonces la convergencia de esta serie en el espacio $L_2[-\pi, \pi]$ hacia

*) El índice en los productos escalares y semiescalares indica en qué espacios se toman los productos en consideración.

el elemento F significa que

$$\lim_{n\to\infty} \|F - S_n(x)\|_{L_2} = 0. \tag{58.54}$$

Si ahora, $f \in F$, se tiene (véase (57.42))

$$\|F - S_n(x)\|_{L_2} = \|f(x) - S_n(x)\|_{RL_2} = \sqrt{\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx}, \tag{58.55}$$

donde $\|f(x) - S_n(x)\|_{RL_2}$ es la seminorma de la función $f(x) - S_n(x)$ en el espacio $RL_2[-\pi, \pi]$, lo que tiene sentido, pues $f(x) - S_n(x) \in F - S_n(x)$. De (58.54) y (58.55) se desprende que

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = \lim_{n\to\infty} \|f(x) - S_n(x)\|_{RL_2} = 0,$$

es decir, la igualdad (58.50) queda demostrada.

Luego, como, en virtud de la misma fórmula (57.42), tienen lugar las igualdades

$$\|F\|_{L_2} = \|f\|_{RL_2} = \sqrt{\int_{-\pi}^{\pi} f^2(x) dx}$$

y los coeficientes de Fourier de F y f son iguales, entonces (58.51) se desprende inmediatamente de (58.53).

Para demostrar el corolario 2, observemos que si todos los coeficientes de Fourier de la función $f \in RL_2[-\pi, \pi]$ según el sistema trigonométrico son nulos, entonces de la igualdad de Parseval (58.51) proviene que

$$\|f\|^2_{RL_2} = \int_{-\pi}^{\pi} f^2(x) dx = 0,$$

y esto significa, de acuerdo con la definición 38 del p. 57.10 referente a las funciones equivalentes, que

$$f \sim 0.$$

Así, fijemos nuestra atención en que si los coeficientes de Fourier de una función de cuadrado integrable son todos iguales a cero, dicha función no es necesariamente cero idéntico, sino que es equivalente a él.

Ambos corolarios están demostrados.

De la igualdad de Parseval (58.51) se infiere una vez más (independientemente del teorema 2 en el p. 55.2) que los coeficientes de Fourier de la función $f(x)$ tienden a cero (pues el término general de la serie convergente (58.51) siempre tiende a cero) con la peculiaridad de que esto tiene lugar sólo para las funciones de cuadrado integrable en el segmento $[-\pi, \pi]$. Por cuanto toda función, continua en el segmento $[-\pi, \pi]$, es a la vez una función de cuadrado integrable, para ella es también válida la afirmación del primer corolario del teorema 12; esta función se desarrolla en serie de Fourier que converge a ella en el sentido de la media cuadrática y para ella es lícita la igualdad de Parseval (58.51).

Mientras tanto el segundo corolario para las funciones continuas puede ser considerablemente reforzado. Enunciémoslo en forma de un teorema.

Teorema 13. *Si todos los coeficientes de Fourier de una función continua en el segmento* $[-\pi, \pi]$ *son iguales a cero, la misma función es idénticamente igual a cero.*

Corolario (teorema sobre la unicidad del desarrollo de una función continua en la serie de Fourier). *Si dos funciones continuas tienen iguales coeficientes de Fourier, son idénticamente iguales.*

DEMOSTRACIÓN. Si la función $f(x)$ es continua en el segmento $[-\pi, \pi]$ y todos los coeficientes de Fourier de la misma son nulos, de la igualdad de Parseval (58.51) tenemos $\|f\|_{RL_2} = 0$. Pero la seminorma del espacio $RL_2[-\pi, \pi]$ en el conjunto de funciones continuas interviene como una norma (véase el ejemplo 9 en el p. 57.4), por lo cual $f(x) = 0$ para cualquier $x \in [-\pi, \pi]$.

El corolario se desprende de lo que la diferencia de dos funciones, cuyos coeficientes de Fourier son iguales, tiene coeficientes de Fourier nulos y, por ende, es cero idéntico. □

OBSERVACIÓN 1. Los teoremas 12 y 13 se han formulado con referencia al sistema trigonométrico de funciones. Las afirmaciones similares son válidas, por supuesto, para cualquier sistema ortogonal completo de funciones, es decir, para un sistema que forma una base ortogonal en el espacio $L_2[a, b]$. En particular, las afirmaciones análogas subsisten para los desarrollos de las funciones según polinomios de Legendre (véase el ejemplo 2 en el p. 58.3) en el espacio $L_2[-1, 1]$. Por ejemplo, si todos los coeficientes de Fourier de una función continua en el segmento $[-1, 1]$ según el sistema de polinomios de Legendre son nulos, dicha función será igual a cero en todo punto del segmento $[-1, 1]$. Las demostraciones de las afirmaciones semejantes pueden efectuarse siguiendo el mismo esquema que se ha usado anteriormente.

OBSERVACIÓN 2. El hecho esencial y fundamental que ha permitido demostrar el teorema 12 consiste en la completitud del sistema trigonométrico en el espacio $L_2[-\pi, \pi]$, la cual, a su vez, se basa sobre la posibilidad de aproximar en el segmento $[-\pi, \pi]$, con cualquier grado de precisión en el sentido de la media cuadrática, toda función de cuadrado integrable en dicho segmento por una función continua de período 2π (véase el lema 16 del p. 57.10). Entre tanto, el empleo de la teoría general sobre el desarrollo según sistemas ortogonales en el espacio de Hilbert tenía, de hecho, sólo carácter terminológico y permitía realizar y anotar los razonamientos en una forma más breve e ilustrativa. A título de ejemplo de un concepto que es muy cómodo, al considerar los problemas estudiados, indiquemos, ante todo, el de espacio lineal normalizado (en particular, del espacio prehilbertiano) y, por consiguiente, el concepto de norma. La introducción de estos conceptos permitió exponer la teoría de desarrollos según los sistemas ortonormalizados independientemente de su forma concreta. Estos conceptos tienen también las más diversas aplicaciones en diferentes apartados de las matemáticas.

En conclusión, haciendo uso de los resultados obtenidos, demostremos un teorema más.

Teorema 14. *Supongamos que la función f es continua en el segmento* $[-\pi, \pi]$. *Si su serie de Fourier converge uniformemente en el segmento* $[-\pi, \pi]$, *la suma de dicha serie será igual a la función f.*

DEMOSTRACIÓN. Sea

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx$$

y

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx$$

la suma de la serie de Fourier de la función f.

Ante todo, la función $S(x)$, como una suma de la serie convergente de las funciones continuas, es también continua. Luego, en virtud del teorema 1 del p. 55.1, los números $a_0, a_n, b_n, n = 1, 2, \ldots$, constituyen los coeficientes de Fourier de la función $S(x)$.

De este modo, dos funciones f y S, continuas en el segmento $[-\pi, \pi]$ tienen coeficientes de Fourier iguales y por esta razón, en virtud de lo dicho anteriormente, coinciden en todos los puntos del segmento $[-\pi, \pi]$: $f(x) = S(x)$, $-\pi \leqslant x \leqslant \pi$. □

58.7°. TRANSFORMACIÓN DE FOURIER DE LAS FUNCIONES INTEGRABLES EN CUADRADO. TEOREMA DE PLANCHEREL

Si el cuadrado de la función f es integrable en todo el eje real, la propia función, en el caso general, no es absolutamente integrable en todo el eje, lo que muestra el ejemplo de la función

$$f(x) = \frac{1}{\sqrt{1 + x^2}}.$$

Por eso, en virtud de la teoría de transformación de Fourier expuesta en el § 56, no puede afirmarse la existencia de la transformación de Fourier para las funciones del espacio $L_2(-\infty, \infty)$. Mostremos que en este caso se puede definir la transformación de Fourier en cierto sentido generalizado. Previamente nos detendremos en la definición del espacio $L_2(-\infty, \infty)$ para las funciones de valores complejos.

Sean f y g dos funciones continuas de cuadrado integrable del módulo en todo el eje que toman, por regla general, los valores complejos. Su producto escalar se determina en este caso según la fórmula

$$(f, g) = \int_{-\infty}^{+\infty} f(x)\overline{g(x)}dx.$$

Se comprueba con facilidad que todas las propiedades, de las cuales debe disponer un producto escalar en un espacio lineal complejo (véase el p. 57.7), en el caso dado se cumplen.

El espacio $L_2(-\infty, \infty)$, que se considera en este punto, se definirá como complemento de un espacio prehilbertiano de funciones continuas de valores

complejos con cuadrado del módulo integrable en todo el eje y producto escalar indicado (compárese con el teorema 6, p. 57.10). Con $\|f\|$ se designa en este párrafo la norma del elemento $f \in L_2(-\infty, +\infty)$, es decir,

$$\|f\| = \sqrt{(f, f)},$$

y asimismo la seminorma

$$\|f\| = \sqrt{\int_{+\infty}^{+\infty} f(x)\overline{f(x)}dx}$$

para las funciones f de cuadrado del módulo integrable en todo el eje. Anteriormente se ha asumido sin demostración, para el caso de las funciones reales (vease el p. 57.10), que todo elemento del espacio L_2 puede considerarse como una clase de funciones. La afirmación análoga es válida también para el espacio L_2 de funciones de valores complejos, con la particularidad de que la seminorma $\|f\|$ de las funciones f coincide con la norma del elemento del espacio L_2, al cual pertenece (en el sentido análogo al indicado en el p. 57.10) la función f. No nos detendremos en la demostración de estos hechos y no los utilizaremos en adelante.

Una función de valores complejos $f(x) = \varphi(x) + i\psi(x)$, donde $\varphi(x)$ y $\psi(x)$ son funciones reales, $-\infty < x < +\infty$, se llamará función escalonada finita, si lo son las funciones $\varphi(x)$ y $\psi(x)$ (véase la definición 7 en el p. 55.2). En lo que sigue las funciones escalonadas finitas se llamarán, para abreviar, simplemente *escalonadas*.

Cualesquiera dos funciones escalonadas $\varphi(x)$ y $\psi(x)$ pueden ser representadas en forma de una combinación lineal finita de las mismas funciones de un escalón (véase el p. 55.2) que toman los valores 1 y 0. Para ello basta tomar toda clase de intersecciones no vacías de los semiintervalos, donde las funciones $\varphi(x)$ y $\psi(x)$ quedan constantes. Dichas intersecciones son también semiintervalos $[x_{k-1}, x_k)$, $k = 1, 2, \ldots, n$, donde quedan constantes simultáneamente las funciones $\varphi(x)$ y $\psi(x)$. Por eso, si

$$\omega_k(x) = \begin{cases} 1, & \text{si } x_{k-1} \leqslant x < x_k, \\ 0, & \text{si } x < x_{k-1} \text{ ó } x \geqslant x_k, \end{cases}$$
$$k = 1, 2, \ldots, n,$$

son las correspondientes funciones de un escalón, existen tales números reales λ_k, μ_k $= 1, 2, \ldots, n$, que

$$\varphi(x) = \sum_{k=1}^{n} \lambda_k \omega_k(x), \quad \psi(x) = \sum_{k=1}^{n} \mu_k \omega_k(x).$$

De aquí se deduce que toda función de un escalón de valores complejos $f(x) = \varphi(x) + i\psi(x)$ es representable en forma de

$$f(x) = \sum_{k=1}^{n} \zeta_k \omega_k(x), \tag{58.56}$$

donde $\zeta_k = \lambda_k + i\mu_k$, $k = 1, 2, \ldots, n$, son números complejos.

Lema 6. *Sea f una función escalonada de valores complejos y sea $F[f]$ su transformación de Fourier, entonces*

$$\|F[f]\| = \|f\|.$$

DEMOSTRACIÓN. Si la función f viene dada por la fórmula (58.56), se tiene

$$\|f\|^2 = \int_{-\infty}^{+\infty} f(x)\overline{f(x)}dx =$$

$$= \sum_{j,k=1}^{n} \zeta_j \overline{\zeta}_k \int_{-\infty}^{+\infty} \omega_j(x)\overline{\omega_k(x)}dx = \sum_{k=1}^{n} |\zeta_k|^2 (x_k - x_{k-1}). \qquad (58.57)$$

Sea ahora $0 < \eta < +\infty$; entonces

$$\int_{-\eta}^{\eta} F[f]\overline{F[f]}dy = \frac{1}{2\pi} \int_{-\eta}^{\eta} dy \int_{-\infty}^{+\infty} f(x)e^{-ixy}dx \int_{-\infty}^{+\infty} \overline{f(\xi)}e^{i\xi y}d\xi =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)\overline{f(\xi)}dx\,d\xi \int_{-\eta}^{\eta} e^{iy(\xi - x)}dy =$$

$$= \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)\overline{f(\xi)} \frac{\operatorname{sen}\eta(\xi - x)}{\xi - x} dx\,d\xi. \qquad (58.58)$$

Todas las transformaciones son verídicas aquí, puesto que todas las integrales se calculan, de hecho, dentro de los límites finitos.

Por cuanto las partes real e imaginaria de la función $f(x)$ satisfacen las condiciones del teorema sobre la representación de funciones mediante la integral de Fourier (véase el teorema 1 en el p. 56.1), resulta, pues, que para todo x, a excepción de $x = x_k$, $k = 1, 2, \ldots, n$, se tiene (véase la demostración del teorema citado)

$$\lim_{\eta \to +\infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \overline{f(\xi)} \frac{\operatorname{sen}\eta(\xi - x)}{\xi - x} d\xi = \overline{f(x)}.$$

Se pone de manifiesto que en virtud de lo observado y teniendo presentes nuestras suposiciones acerca de la última integral (58.58), podemos pasar al límite bajo el signo de la integral exterior para $\eta \to +\infty$. No obstante, el teorema correspondiente no fue demostrado en el presente curso, a consecuencia de lo cual nos vemos obligados a realizar ciertos cálculos complementarios. Sustituyendo (58.56) en (58.58), obtenemos

$$\int_{-\eta}^{\eta} F[f]\overline{F[f]}dy = \frac{1}{\pi} \sum_{j,k=1}^{n} \zeta_j \overline{\zeta}_k \int_{x_{j-1}}^{x_j} dx \int_{x_{k-1}}^{x_k} \frac{\operatorname{sen}\eta(\xi - x)}{\xi - x} d\xi =$$

$$= \frac{1}{\pi} \sum_{j,k=1}^{n} \zeta_j \overline{\zeta}_k \int_{x_{j-1}}^{x_j} dx \int_{\eta(x_{k-1} - x)}^{\eta(x_k - x)} \frac{\operatorname{sen} t}{t} dt. \qquad (58.59)$$

Analicemos el comportamiento de cada sumando de la suma obtenida cuando $\eta \to +\infty$. Si $j = k$, entonces, cambiando el orden de integración (fig. 242) y efec-

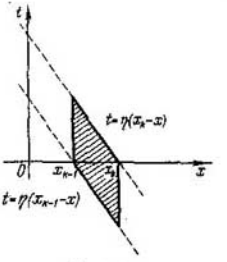


Fig. 242

tuando la integración respecto de la variable x, obtendremos:

$$\frac{1}{\pi}\int_{x_{k-1}}^{x_k} dx \int_{\eta(x_{k-1}-x)}^{\eta(x_k-x)} \frac{\operatorname{sen} t}{t}\,dt = \cdot$$

$$= \frac{1}{\pi}\left[\int_0^{\eta(x_k-x_{k-1})} \left(x_k - x_{k-1} - \frac{t}{\eta}\right)\frac{\operatorname{sen} t}{t}\,dt + \right.$$

$$\left. + \int_{-\eta(x_k-x_{k-1})}^{0} \left(x_k - x_{k-1} + \frac{t}{\eta}\right)\frac{\operatorname{sen} t}{t}\,dt\right] =$$

$$= \frac{2}{\pi}(x_k - x_{k-1}) \int_0^{\eta(x_k-x_{k-1})} \frac{\operatorname{sen} t}{t}\,dt + \frac{2}{\pi\eta}[1 - \cos\eta(x_k - x_{k-1})].$$

Como

$$\int_0^{+\infty} \frac{\operatorname{sen} t}{t}\,dt = \frac{\pi}{2}$$

(véase el p. 54.4), se tiene

$$\lim_{\eta\to+\infty} \frac{2}{\pi}(x_k - x_{k-1}) \int_0^{\eta(x_k-x_{k-1})} \frac{\operatorname{sen} t}{t}\,dt = x_k - x_{k-1}.$$

Luego, evidentemente,

$$\lim_{\alpha\to+\infty} \frac{2}{\pi\eta}[1 - \cos\eta(x_k - x_{k-1})] = 0,$$

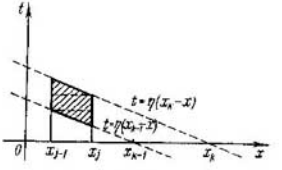


Fig. 243

por lo cual

$$\lim_{\eta \to +\infty} \frac{1}{\pi} \int_{x_{k-1}}^{x_k} dx \int_{\eta(x_{k-1}-x)}^{\eta(x_k-x)} \frac{\operatorname{sen} t}{t}\, dt = x_k - x_{k-1}, \quad k = 1, 2, \ldots, n.$$

Mostremos ahora que para $j \neq k$

$$\lim_{\eta \to +\infty} \int_{x_{j-1}}^{x_j} dx \int_{\eta(x_{k-1}-x)}^{\eta(x_k-x)} \frac{\operatorname{sen} t}{t}\, dt = 0.$$

Sea, para concretar, $x_{j-1} < x_j \leqslant x_{k-1} < x_k$. Para otras disposiciones de los semiintervalos $[x_{j-1}, x_j)$ y $[x_{k-1}, x_k)$, donde las funciones quedan constantes, la demostración es análoga. Cambiando de nuevo el orden de integración y realizando la integración respecto de x (fig. 243) obtendremos, recurriendo a los razonanientos semejantes:

$$\int_{x_{j-1}}^{x_j} dx \int_{\eta(x_{k-1}-x)}^{\eta(x_k-x)} \frac{\operatorname{sen} t}{t}\, dt = \int_{\eta(x_{k-1}-x_j)}^{\eta(x_k-x_j)} \left(x_k - x_{j-1} - \frac{t}{\eta}\right) \frac{\operatorname{sen} t}{t}\, dt +$$

$$+ \int_{\eta(x_{k-1}-x_{j-1})}^{\eta(x_k-x_j)} (x_j - x_{j-1}) \frac{\operatorname{sen} t}{t}\, dt +$$

$$+ \int_{\eta(x_{k-1}-x_j)}^{\eta(x_{k-1}-x_{j-1})} \left(x_j - x_{k-1} + \frac{t}{\eta}\right) \frac{\operatorname{sen} t}{t}\, dt \to 0 \quad \text{para} \quad \eta \to \infty.$$

Ahora, de (58.59) tenemos

$$\|F[f]\|^2 = \int_{-\infty}^{+\infty} F[f]\overline{F[f]}\, dx = \lim_{\eta \to +\infty} \int_{-\eta}^{\eta} F[f]\overline{F[f]}\, dy =$$

$$= \sum_{k=1}^{n} |\xi_k|^2 (x_k - x_{k-1}) = \|f\|^2. \quad \square$$

Lema 7. *Sea f una función de valores complejos, continua en el segmento* $[a, b]$ *e igual a cero fuera de dicho segmento; existe una sucesión de tales funciones escalonadas* φ_n, $n = 1, 2, \ldots$, *que*

$$\lim_{n \to \infty} \|\varphi - \varphi_n\| = 0.$$

DEMOSTRACIÓN. Para las funciones reales dicha afirmación se desprende del lema 14, p. 57.10. Supongamos ahora que $\varphi = u + iv$ es una función de valores complejos continua en el segmento $[a, b]$; en este caso las funciones reales u y v son también continuas en el segmento $[a, b]$. Por eso, existen tales sucesiones de funciones escalonadas $\{u_n\}$ y $\{v_n\}$ que $\|u - u_n\| \to 0$ y $\|v - v_n\| \to 0$ cuando $n \to \infty$. Si es que $\varphi_n = u_n + iv_n$, entonces $\|\varphi - \varphi_n\| \leqslant \|u - u_n\| + \|v - v_n\|$, de donde $\|\varphi - \varphi_n\| \to 0$ cuando $n \to \infty$. □

Lema 8. *Supongamos que una función de valores complejos* φ *es continua en el segmento* $[a, b]$ *y es igual a cero fuera de él, en este caso*

$$\|F[\varphi]\| = \|\varphi\|.$$

DEMOSTRACIÓN. Sea φ_n una sucesión de funciones escalonadas de tal género que

$$\lim_{n \to \infty} \|\varphi - \varphi_n\| = 0$$

(véase el lema 7), entonces, en virtud de la continuidad de la norma,

$$\lim_{n \to \infty} \|\varphi_n\| = \|\varphi\|. \tag{58.60}$$

Entre tanto, de la desigualdad de Cauchy — Buniakovski obtenemos

$$\int_a^b |\varphi_n(x) - \varphi(x)|\,dx \leqslant \left(\int_a^b dx\right)^{1/2} \left(\int_a^b |\varphi_n(x) - \varphi(x)|^2 dx\right)^{1/2} = (b - a)^{1/2} \left(\int_a^b |\varphi_n(x) - \varphi(x)|^2 dx\right)^{1/2}$$

y, por consiguiente,

$$\lim_{n \to \infty} \int_a^b |\varphi_n(x) - \varphi(x)|\,dx = 0,$$

es decir, la sucesión $\{\varphi_n\}$ converge en media hacia la función φ y converge en el sentido de L_1. Por esta razón, si

$$\psi = F[\varphi], \quad \psi_n = F[\varphi_n], \quad n = 1, 2, \ldots,$$

entonces la sucesión de las funciones continuas $\{\psi_n\}$ (véase el corolario del lema 4 en el p. 56.7) converge uniformemente hacia la función ψ, la cual es, debido a esto, continua en todo el eje numérico. Además, en virtud del lema 6,

$$\|\psi_n\| = \|\varphi_n\|. \tag{58.61}$$

De aquí proviene, en particular, que las funciones continuas ψ_n son funciones de cuadrado integrable del módulo, es decir, pertenecen al espacio $L_2(-\infty, +\infty)$. Ahora, las funciones ψ_n, $n = 1, 2, \ldots$, forman una sucesión fundamental en el espacio $L_2(-\infty, +\infty)$. Esto se deduce de lo que la sucesión $\{\varphi_n\}$ converge en media

en el sentido de L_2 y también de la igualdad

$$\int_{-\infty}^{+\infty} |\psi_n(y) - \psi_m(y)|^2 dy = \int_{-\infty}^{+\infty} |\varphi_n(y) - \varphi_m(y)|^2 dy,$$

la cual también proviene del lema 6, pues la diferencia entre las funciones escalonadas es asimismo una función escalonada.

Mostremos que la sucesión $\{\psi_n\}$ converge hacia la función ψ también en el espacio L_2. En efecto, siendo $\varepsilon > 0$ fijo, existe, por ser la sucesión $\{\psi_n\}$ fundamental, tal número n_ε que para todos los $n \geqslant n_\varepsilon$ y $m \geqslant n_\varepsilon$ se cumple la desigualdad

$$\|\psi_n - \psi_m\|^2 = \int_{-\infty}^{+\infty} |\psi_n(y) - \psi_m(y)|^2 dy < \varepsilon.$$

Con mayor razón, para cualquier número $c > 0$ tendremos

$$\int_{-c}^{c} |\psi_n(y) - \psi_m(y)|^2 dy < \varepsilon. \tag{58.62}$$

Cuando n y c son fijos y $m \to \infty$, la expresión subintegral en (58.62) converge uniformemente hacia la función $|\psi_n(y) - \psi(y)|^2$. Por ello, en la desigualdad (58.62) podemos pasar al límite bajo el signo de la integral para $m \to \infty$. De resultas tendremos

$$\int_{-c}^{c} |\psi_n(y) - \psi(y)|^2 dy \leqslant \varepsilon.$$

Ahora, haciendo tender c hacia $+\infty$, obtendremos que para $n \geqslant n_\varepsilon$ se verifica la desigualdad

$$\int_{-\infty}^{+\infty} |\psi_n(y) - \psi(y)|^2 dy \leqslant \varepsilon, \tag{58.63}$$

lo que precisamente significa la convergencia en media, en el sentido de L_2, de la sucesión $\{\psi_n\}$ hacia la función ψ.

De lo demostrado se deduce también que $\psi \in L_2(-\infty, +\infty)$. Efectivamente, en virtud de (58.61) y (58.63),

$$\|\psi\| \leqslant \|\psi - \psi_n\| + \|\psi_n\| < +\infty.$$

Por fin, de la desigualdad (57.18) obtenemos, teniendo presente que $\lim_{n\to\infty} \|\psi_n - \psi\| = 0$,

$$\lim_{n\to\infty} \|\psi_n\| = \|\psi\|. \tag{58.64}$$

De (58.60), (58.61) y (58.64) se infiere que

$$\|\psi\| = \|\varphi\|. \quad \square$$

Teorema 15 (de Plancherel*)). *Supongamos que φ es una función continua de cuadrado integrable del módulo en todo el eje numérico y sea*

*) M. Plancherel (1885 — 1967), matemático suizo.

$$\psi_M(y) = \frac{1}{\sqrt{2\pi}} \int_{-M}^{M} \varphi(x) e^{-ixy} dx, \quad M > 0.$$

Entonces:

1) *la función $\psi_M(y)$ es también continua y de cuadrado integrable en todo el eje numérico;*

2) *cuando $M \to +\infty$, las funciones ψ_M convergen en el espacio $L_2(-\infty, +\infty)$ hacia cierto elemento $\psi \in L_2(-\infty, +\infty)$ y*

3) $\|\varphi\| = \|\psi\|$.

DEMOSTRACIÓN. Si

$$\varphi_M(x) = \begin{cases} \varphi(x), & \text{si } x \in [-M, M], \\ 0, & \text{si } x \notin [-M, M], \end{cases}$$

será obvio que

$$\psi_M = F[\varphi_M],$$

$$\lim_{M \to +\infty} \varphi_M = \varphi \text{ en } L_2(-\infty, +\infty), \tag{58.65}$$

$$\lim_{M \to +\infty} \|\varphi_M\| = \|\varphi\|. \tag{58.66}$$

De conformidad con el lema 8,

$$\|\psi_M\| = \|\varphi_M\|, \quad M > 0, \tag{58.67}$$

$$\|\psi_{M_1} - \psi_{M_2}\| = \|\varphi_{M_1} - \varphi_{M_2}\|, \quad M_1 > 0, \quad M_2 > 0. \tag{58.68}$$

De (58.65) y (58.68) proviene, en virtud de que el espacio $L_2(-\infty, +\infty)$ es completo, que existe un límite (¿por qué?)

$$\lim_{M \to +\infty} \psi_M = \psi \text{ en } L_2(-\infty, +\infty).$$

Por ser continua la norma,

$$\lim_{M \to +\infty} \|\psi_M\| = \|\psi\| \tag{58.69}$$

de (58.66), (58.67) y (58.69) tenemos

$$\|\psi\| = \|\varphi\|. \ \square$$

El elemento $\psi \in L_2(-\infty, +\infty)$, obtenido en el transcurso de la demostración, lo llamaremos también *transformación de Fourier* de la función continua dada $\varphi \in L_2(-\infty, +\infty)$ y anotaremos

$$\psi = F[\varphi]. \tag{58.70}$$

Esta notación es natural, puesto que si la función φ es, además, absolutamente integrable, entonces $\lim_{M \to +\infty} \psi_M$ coincide con la transformación de Fourier ordinaria. En efecto, en este caso

$$\lim_{M \to +\infty} \int_{-\infty}^{+\infty} |\varphi_M(x) - \varphi(x)| dx = 0.$$

Por consiguiente, las funciones $\psi_M = F[\varphi_M]$ convergen uniformemente, cuando $M \to \infty$, hacia la transformación de Fourier $F[\varphi]$ de la función φ. Como vimos, ψ_M converge en media en el sentido de L_2 hacia la función ψ; de aquí no es difícil convencerse de que $\psi = F[\varphi]$ (compárese el razonamiento análogo en la demostración del lema 8).

La transformación de Fourier (58.70) está definida por ahora sólo para aquellos elementos $\varphi \in L_2(-\infty, +\infty)$, que representan en sí funciones continuas de cuadrado integrable, sin embargo por su continuidad puede ser extendida a todo el espacio $L_2(-\infty, +\infty)$. En efecto, sea φ un elemento arbitrario del espacio $L_2(-\infty, +\infty)$. De acuerdo con la definición del espacio $L_2(-\infty, +\infty)$, el conjunto de funciones continuas es denso en él. Por consiguiente, existe una sucesión de funciones continuas

$$\varphi_n \in L_2(-\infty, +\infty), \quad n = 1, 2, \ldots,$$

tal que $\lim_{n\to\infty} \varphi_n = \varphi$, es decir, $\lim_{n\to\infty} \|\varphi_n - \varphi\| = 0$.

Sea $F[\varphi_n] = \psi_n$, $n = 1, 2, \ldots$. En virtud del teorema Plancherel,

$$\|\psi_n - \psi_m\| = \|\varphi_n - \varphi_m\|, \quad n, m = 1, 2, \ldots,$$

por lo cual la sucesión $\{\psi_n\}$ es fundamental en L_2 y, por ende, es convergente. Sea $\psi = \lim_{n\to\infty} \psi_n$. Por definición suponemos

$$\psi = F[\varphi]. \tag{58.71}$$

Si $\varphi_n^* \in L_2(-\infty, +\infty)$, $n = 1, 2, \ldots$, es alguna otra sucesión de funciones continuas que en $L_2(-\infty, +\infty)$ converge al elemento φ, y si $\psi_n^* = F[\varphi_n^*]$, entonces de la igualdad

$$\|\varphi_n - \varphi_n^*\| = \|\psi_n - \psi_n^*\|$$

tenemos $\lim_{n\to\infty} \psi_n^* = \psi$. De este modo, la definición (58.71) no depende de cómo se elige la sucesión de funciones continuas que converja hacia el elemento φ.

Para cualquier $\varphi \in L_2(-\infty, +\infty)$ resulta válida la igualdad

$$\|F[\varphi]\| = \|\varphi\|,$$

lo cual proviene inmediatamente de que esta igualdad tiene lugar para las funciones continuas $\varphi \in L_2(-\infty, +\infty)$ y de la continuidad de la norma.

Luego, es fácil comprobar que la transformación de Fourier F es lineal en $L_2(-\infty, +\infty)$, es decir,

$$F[\lambda_1\varphi_1 + \lambda_2\varphi_2] = \lambda_1 F[\varphi_1] + \lambda_2 F[\varphi_2]$$

para cualesquiera φ_1 y φ_2 de $L_2(-\infty, +\infty)$ y cualesquiera números λ_1 y λ_2. Esto es cierto para las funciones escalonadas. Ellas forman en $L_2(-\infty, +\infty)$ un conjunto denso. De aquí, la igualdad citada se obtiene pasando al límite para cualesquiera elementos del espacio $L_2(-\infty, +\infty)$.

Por fin, la transformación de Fourier aplica el espacio $L_2(-\infty, +\infty)$ *sobre sí mismo*, es decir, cualquiera que sea el elemento $\psi \in L_2(-\infty, +\infty)$, existe un ele-

mento $\varphi \in L_2(-\infty, +\infty)$ tal que $F[\varphi] = \psi$. Con el fin de probarlo, se debe definir, empleando el mismo método que se ha utilizado para la transformación de Fourier, en el espacio $L_2(-\infty, +\infty)$ la transformación de Fourier inversa F^{-1} y mostrar que para todo elemento $\psi \in L_2(-\infty, +\infty)$ se verifica la igualdad $\|F^{-1}[\psi]\| = \|\psi\|$. A continuación, se puede mostrar que

$$F[F^{-1}[\psi]] = \psi \quad \text{y} \quad F^{-1}[F[\psi]] = \psi$$

para cualquier $\psi \in L_2(-\infty, +\infty)$, partiendo de que esto es cierto sobre un conjunto de funciones escalonadas que forman en $L_2(-\infty, +\infty)$ un conjunto denso. Si, ahora, tomamos, para el elemento $\psi \in L_2(-\infty, +\infty)$, un elemento $\varphi = F^{-1}[\psi]$, entonces obtendremos $F[\varphi] = \psi$, lo que precisamente significa que la transformación F aplica todo el espacio $L_2(-\infty, +\infty)$ sobre sí mismo.

Al resumir todo lo dicho, llegamos al siguiente teorema.

Teorema 16 (de Plancherel). *La transformación de Fourier F es lineal y aplica biunívocamente el espacio $L_2(-\infty, +\infty)$ sobre sí mismo, con la particularidad de que para cualquier elemento $\varphi \in L_2(-\infty, +\infty)$ se verifica la igualdad*

$$\|F[\varphi]\| = \|\varphi\|.$$

§ 59. FUNCIONES GENERALIZADAS

59.1. RAZONAMIENTOS GENERALES

Se considerará en este párrafo una generalización del concepto clásico de una función, a saber, el concepto de función generalizada. Ha surgido en el proceso de resolución de ciertos problemas físicos y en los últimos años se ha arraigado en las matemáticas rápida y sólidamente. Con ayuda de este concepto la transformación de Fourier puede extenderse a una clase de funciones más amplia que las funciones absolutamente integrables o de cuadrado integrable. Permite formular en el lenguaje matemático tales nociones idealizadas como, por ejemplo, densidad de una carga puntual, densidad de un punto material, impulso instantáneo, etc.

Explicaremos esto más detalladamente. Al estudiar fenómenos físicos sirviéndonos del aparato matemático, hemos de recurrir inevitablemente a ciertas abstracciones matemáticas, en particular, emplear la noción de punto. Hablamos, por ejemplo, de una masa concentrada en el punto dado de un espacio, de una fuerza aplicada en el momento dado de tiempo (es decir, en el punto dado del eje a lo largo del cual se calcula el tiempo), de una fuente puntual de tal o cual campo físico, etc. Esto resulta ser cómodo al emplear el aparato matemático, aunque en tal caso reproducimos una situación que no es del todo exacta: cualquier masa tiene un volumen bien determinado, toda fuerza actúa durante un determinado lapso, cualquier fuente de un campo tiene dimensiones determinadas, etc. Resulta pues que, para este enfoque del estudio de los fenómentos físicos, los métodos clásicos de las matemáticas son insuficientes. A veces nos vemos obligados a introducir nuevos conceptos matemáticos, crear un aparato matemático nuevo.

Examinemos, a título de ejemplo, la acción de una fuerza "instantánea". Supongamos que en el instante $t = 0$ un cuerpo de masa $m \neq 0$ ha experimentado la acción de una fuerza que le comunicó la velocidad $v \neq 0$, después de lo cual la acción de la fuerza se dio por terminada. Al designar mediante $F(t)$ la fuerza que actúa contra el cuerpo al instante de tiempo t, obtenemos $F(t) = 0$ cuando $t \neq 0$. Trataremos de determinar cuál es la fuerza $F(t)$ cuando $t = 0$. Según la segunda ley de Newton, la fuerza es igual a la velocidad de variación de la cantidad de movimiento respecto del tiempo

$$F(t) = \frac{d(mv)}{dt}$$

y, por consiguiente, para cualquier momento de tiempo τ, $0 < \tau < +\infty$, se tiene

$$\int_{-\infty}^{\tau} F(t)\,dt = mv. \tag{59.1}$$

Como límite inferior de integración se ha elegido $-\infty$; por supuesto, podemos tomar en lugar de $-\infty$ cualquier número $a < 0$, por cuanto hasta el momento de tiempo $t = 0$ el cuerpo se encontraba en reposo.

Fijemos nuestra atención en que desde el punto de vista de la matemática clásica, es decir, desde el punto de vista de aquel concepto de integral que hemos estudiado, la igualdad (59.1) está privada de sentido: la función $F(x)$ es igual a cero en todos los puntos, salvo en $t = 0$, por eso la integral que figura en el primer miembro de la fórmula (59.1) y que se considera como impropia es igual a cero, mientras que el segundo miembro de esta igualdad no es nulo. Por otra parte, partiendo de los razonamientos físicos, es natural esperar que la igualdad escrita tiene un sentido determinado. Esta contradicción significa que nos encontramos fuera de las posibilidades de emplear el aparato matemático conocido y han de introducirse algunas nuevas nociones matemáticas.

Supongamos, para simplificar, que la cantidad de movimiento adquirido por el cuerpo es igual a la unidad, es decir, $mv = 1$. En este caso la fuerza $F(t)$ que actúa contra el cuerpo se designará mediante $\delta(t)$, por lo cual, la fórmula (59.1) tendrá por expresión

$$\int_{-\infty}^{\tau} \delta(t)\,dt = 1, \quad \tau > 0. \tag{59.2}$$

La función $\delta(t)$ se llama corrientemente función delta (función δ) o función de Dirac *).

Con el fin de examinar detenidamente la cuestión, supongamos que el cuerpo está accionado no por una fuerza instantánea, sino por cierta fuerza constante que actúa contra el cuerpo durante el lapso de tiempo de $-\varepsilon$ hasta 0 $(\varepsilon > 0)$; esta última se designará con $\delta_\varepsilon(t)$. Supongamos también que esta fuerza comunica al cuerpo la misma cantidad de movimiento, igual a la unidad. En otras palabras, distribuyamos

*) P. Dirac (n. 1902), físico inglés.

la fuerza buscada $\delta(t)$ por un intervalo de tiempo de longitud ε. Hallemos la fuerza $\delta_\varepsilon(t)$.

De acuerdo con el principio de conservación de la cantidad de movimiento, para cualquier tiempo $\tau \geqslant 0$ tenemos

$$\int_{-\infty}^{\tau} \delta_\varepsilon(t)\,dt = 1.$$

Por cuanto la fuerza $\delta_\varepsilon(t)$ es nula fuera del segmento $[-\varepsilon, 0]$ y constante dentro del segmento citado, se tiene

$$1 = \int_{-\infty}^{\tau} \delta_\varepsilon(t)\,dt = \int_{-\varepsilon}^{0} \delta_\varepsilon(t)\,dt = \varepsilon\delta_\varepsilon(t), \quad -\varepsilon \leqslant t \leqslant 0.$$

Por ello

$$\delta_\varepsilon(t) = \begin{cases} \dfrac{1}{\varepsilon}, & \text{si} \quad -\varepsilon \leqslant t \leqslant 0 \\ 0, & \text{si} \quad t < -\varepsilon \quad \text{o bien} \quad t > 0. \end{cases} \tag{59.3}$$

Es natural suponer que la fuerza instantánea $\delta(t)$ se obtiene de la "fuerza distribución" $\delta_\varepsilon(t)$, pasando al límite para $\varepsilon \to 0$, es decir,

$$\delta(t) = \lim_{\varepsilon \to 0} \delta_\varepsilon(t),$$

entonces

$$\delta(t) = \begin{cases} +\infty, & \text{si} \quad t = 0, \\ 0, & \text{si} \quad t \neq 0. \end{cases} \tag{59.4}$$

Esta formula no permite obtener la (59.2), haciendo uso de las definiciones conocidas de la integral (propia o impropia). El hecho de que la función es nula en todos los puntos, a excepción de uno, donde ella es igual al infinito, y a la vez otro hecho, consistente en que la integral de dicha función es igual a la unidad, contradicen uno al otro dentro de los marcos de la matemática que hoy día se llama clásica. Esto nos lleva a la conclusión sobre la necesidad de introducir un concepto nuevo, el de "integral" (59.2).

Desde el punto de vista físico es natural considerar que la cantidad de movimiento comunicada al cuerpo por la fuerza instantánea $\delta(t)$, es decir, la integral (59.2), es un límite de la cantidad de movimiento comunicada al cuerpo por las fuerzas $\delta_\varepsilon(t)$ distribuidas en tiempo, cuando el lapso de su accionamiento tiende a cero, es decir, cuando $\varepsilon \to 0$. Por eso pongamos, por definición,

$$\int_{-\infty}^{\tau} \delta(t)\,dt = \lim_{\varepsilon \to 0} \int_{-\infty}^{\tau} \delta_\varepsilon(t)\,dt, \quad \tau > 0.$$

De aquí, en virtud de la igualdad $\int_{-\infty}^{\tau} \delta_\varepsilon(t)\,dt = 1$, $\tau > 0$ para todo $\varepsilon > 0$, proviene precisamente la igualdad (59.2).

De este modo, cuando se dice que la integral (59.2) de una función delta es igual a la unidad, esta integral debe entenderse como límite de las correspondientes integrales ordinarias de las funciones δ_ε cuando $\varepsilon \to +0$.

Resulta útil dar, de un modo análogo, la definición de las "integrales" más generales, a saber, de las integrales del tipo

$$\int_{-\infty}^{\tau} \delta(t) f(t)\, dt, \quad -\infty < \tau < +\infty, \tag{59.5}$$

donde $f(t)$ es una función continua. A saber, definamos el símbolo (59.5) mediante la igualdad

$$\int_{-\infty}^{\tau} \delta(t) f(t)\, dt = \lim_{\varepsilon \to 0} \int_{-\infty}^{\tau} \delta_\varepsilon(t) f(t)\, dt. \tag{59.6}$$

Para demostrar que esta definición es correcta, se debe probar que el límite (59.6) siempre existe. Más aún, probemos que

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\tau} \delta_\varepsilon(t) f(t)\, dt = \begin{cases} f(0) & \text{para } \tau \geqslant 0, \\ 0 & \text{para } \tau < 0. \end{cases} \tag{59.7}$$

Sea primero $\tau \geqslant 0$. Al emplear (59.3), obtendremos

$$\left| \int_{-\infty}^{\tau} \delta_\varepsilon(t) f(t)\, dt - f(0) \right| = \left| \frac{1}{\varepsilon} \int_{-\varepsilon}^{0} f(t)\, dt - \frac{f(0)}{\varepsilon} \int_{-\varepsilon}^{0} dt \right| \leqslant$$

$$\leqslant \frac{1}{\varepsilon} \int_{-\varepsilon}^{0} |f(t) - f(0)|\, dt. \tag{59.8}$$

Por ser la función $f(x)$ continua cuando $x = 0$, para cualquier $\eta > 0$ existe tal $\varepsilon_\eta > 0$ que para todo t que satisfaga la condición $|t| < \varepsilon_\eta$, se cumple la desigualdad

$$|f(t) - f(0)| < \eta.$$

Por ello, para cualquier $\varepsilon < \varepsilon_\eta$, de la desigualdad (59.8) proviene que

$$\left| \int_{-\infty}^{\tau} \delta_\varepsilon(t) f(t)\, dt - f(0) \right| < \frac{\eta}{\varepsilon} \int_{-\varepsilon}^{\tau} dt = \eta.$$

La igualdad (59.7) queda demostrada para $\tau \geqslant 0$. Se demuestra de modo aún más fácil cuando $\tau < 0$.

Así pues, de la definición (59.6) se deduce que para cualquier función continua $f(t)$ es válida la fórmula

$$\int_{-\infty}^{\tau} \delta(t) f(t)\, dt = \begin{cases} f(0) & \text{para } \tau \geqslant 0, \\ 0 & \text{para } \tau < 0. \end{cases} \tag{59.9}$$

La fórmula (59.2) se desprende de aquí cuando $f(t) = 1$.

Si ponemos

$$\theta(t) = \begin{cases} 1 & \text{para } t \geqslant 0, \\ 0 & \text{para } t < 0, \end{cases} \tag{59.10}$$

la fórmula (59.9), para $f(t) = 1$, se escribirá en la forma

$$\theta(\tau) = \int_{-\infty}^{\tau} \delta(t)\, dt. \tag{59.11}$$

La función $\theta(t)$ lleva un nombre especial: se llama *función de Heaviside**). Al calcular la derivada de la función $\theta(t)$, de acuerdo con la definición clásica de derivada, de (59.10) obtenemos

$$\theta'(t) = \begin{cases} \infty & \text{para } t = 0, \\ 0 & \text{para } t \neq 0. \end{cases} \tag{59.12}$$

No sería correcto afirmar a base de lo dicho que $\theta'(t)$ es una función delta, puesto que la función $\delta(t)$ se define no sólo por la función (59.4), pues incluso desde el punto de vista físico está claro que de esta sola fórmula no puede provenir que la fuerza $\delta(t)$ comunica al cuerpo en consideración una cantidad de movimiento que es precisamente igual a la unidad. No obstante, es cómodo poner, por definición,

$$\theta'(t) = \delta(t).$$

Aparte de la igualdad (59.12), esto se justifica por el hecho de que en tal caso queda en vigor la fórmula fundamental del cálculo integral que restablece la función a partir de su derivada, o sea, la fórmula de Newton — Leibniz. En efecto, ahora la fórmula (59.11) puede ser escrita en la forma

$$\theta(\tau) = \int_{-\infty}^{\tau} \theta'(t)\, dt, \quad -\infty < \tau < +\infty$$

(observemos que $\theta(-\infty) = 0$).

Ha de notarse que no hemos dado la definición matemática precisa de la propia función $\delta(t)$ en su calidad de función de un punto (se ha observado anteriormente que la fórmula (59.4) no constituye tal definición); esto es imposible en general, puesto que la función delta es un concepto de distinta naturaleza. Entre tanto, fue definida por nosotros no la función $\delta(t)$, sino la "integral" (59.5). No es por casualidad. Para varios problemas de física resulta característica una circunstancia consis-

*) O. Heaviside (1850—1925), físico inglés.

tente en que las funciones, introducidas con el objeto de describir uno u otro objeto, sólo tienen sentido en la medida de que el sentido físico inmediato lo tienen ciertas integrales de dichas funciones. Las funciones generalizadas surgen, pues, como cierta generalización de una familia de integrales del producto de dos funciones, una de las cuales es fija y la otra puede elegirse arbitrariamente de cierta totalidad.

Así pues, hemos definido un concepto nuevo, el de integral de la función delta (e, incluso, el concepto más, general de integral del producto de una función continua por la función delta). Esto no es una integral ordinaria, es decir, no es un límite de ciertas sumas integrales, sino el límite de las integrales correspondientes, o bien, hablando metafóricamente, "un límite de los límites de sumas integrales". En otras palabras, para definir la integral $\int_{-\infty}^{+\infty} \delta(x)f(x)\,dx$ se debe añadir al paso límite, que da el valor de la integral $\int_{-\infty}^{+\infty} \delta_\varepsilon(x)f(x)\,dx$, un paso límite más para $\varepsilon \to 0$. Aquí se observa una analogía peculiar con la definición de la integral impropia: partiendo de la definición conocida de la integral, obtenemos, con ayuda de un paso límite complementario, el nuevo concepto matemático. Por supuesto, los pasos límite complementarios en estos casos son diferentes, lo que conduce a los distintos conceptos.

Al definir de un modo nuevo el símbolo (59.5), nos encontramos dentro del círculo de las definiciones matemáticas acostumbradas que hacen más amplio el bagaje de conceptos que hemos tratado anteriomente; hemos logrado establecer una propiedad interesante de la función $\delta(t)$ (véase (59.9)): dicha función pone a toda función continua $f(t)$ en correspondencia un número $f(0)$, es decir, $\delta(t)$ puede considerarse como una función definida en el conjunto de todas las funciones continuas. Las aplicaciones cuyos dominios de definición representan en sí ciertos conjuntos de funciones llevan el nombre de *funcionales*. La función delta es precisamente uno de los más simples ejemplos de funcionales. Las funciones generalizadas mencionadas al principio de este punto, se llaman funcionales del tipo determinado (véase el p. 59.2).

Según se vio, las propiedades de la función delta se determinan por las propiedades de las funciones $\delta_\varepsilon(x)$. Si tomamos $\varepsilon = \frac{1}{n}$, $n = 1, 2, \ldots$, se obtendrá una sucesión de funciones la cual, al igual que las sucesiones análogas a la mencionada en un sentido determinado, se llama sucesión en delta (la definición precisa de las sucesiones en delta se dará más abajo: véase el ejercicio 6 en el p. 59.3). Cualquier sucesión en delta puede servir para definir la propiedad (59.9) de la función delta. Cabe notar que ya nos encontramos antes con las sucesiones en delta: a título de ejemplo de tal sucesión sirve la sucesión de los núcleos de Fejer $\Phi_n(n)$, $n = 1, 2, \ldots$. Sin embargo, no fijamos nuestra atención en las sucesiones de esta índole, por cuanto ellas no fueron el objeto de estudio especial y desempeñaban un papel auxiliar.

Pasemos ahora al estudio sistemático de las funciones generalizadas. Algunas de las funciones generalizadas han aparecido primeramente en las obras de P. Dirac y otros físicos en calidad de un método simbólico para describir ciertos fenómenos físicos. Con el fin de emplear estos conceptos a título de método de la investigación

teórica, tuvo que surgir la necesidad de crear la teoría de funciones generalizadas, lo que se ha hecho. La teoría de las funciones generalizadas es ahora un aparato matemático muy útil. Con ayuda de esta teoría se ha logrado resolver toda una serie de problemas que no podían ser resueltos mediante los métodos antiguos. En el presente las funciones generalizadas se utilizan ampliamente tanto en las inverstigaciones aplicadas como en las matemáticas puras.

En los puntos ulteriores de este párrafo expondremos los fundamentos de la teoría general de las funciones generalizadas, elaborada por S. L. Sóbolev y L. Schwartz*).

59.2. ESPACIOS LINEALES CON CONVERGENCIA. FUNCIONALES. ESPACIOS CONJUGADOS

Definición 1. *Sea X cierto conjunto y supongamos que en la totalidad de todas las sucesiones $\{x_n\}$ de sus elementos, $x_n \in X$, se ha distinguido una clase de sucesiones, llamadas convergentes, y a toda sucesión convergente se le ha puesto en correspondencia un elemento $x \in X$, llamado límite de dicha sucesión.*

Si en este caso se cumplen tres condiciones:

1) cada sucesión de elementos del conjunto X puede tener no más de un límite;

2) toda sucesión del tipo $\{x, x, x, \ldots, x, \ldots\}$ es convergente y el elemento x es su límite;

3) toda subsucesión de una sucesión convergente es también convergente y tiene el mismo límite que toda la sucesión,

entonces el conjunto X se denomina espacio con convergencia.

Las condiciones 1, 2 y 3 se llaman *axiomas de Fréchet***).

Si x es el límite de la sucesión $\{x_n\}$, entonces, como siempre, se escribe

$$x = \lim_{n \to \infty} x_n.$$

Definición 2. *Un espacio lineal X se llama espacio lineal con convergencia, si es un espacio con convergencia, respecto de la cual las operaciones de sumación de los elementos del espacio y de su multiplicación por un número son continuas.*

Esto significa que para cualesquiera sucesiones convergentes $\{x_n\}$ e $\{y_n\}$ de elementos de X, que tienen por límites $x \in X$ e $y \in X$, respectivamente, y cualesquiera números λ y μ la sucesión $\{\lambda x_n + \mu y_n\}$ es también convergente y

$$\lim_{n \to \infty} (\lambda x_n + \mu y_n) = \lambda x + \mu y.$$

Además, si $\{\lambda_n\}$ es una sucesión numérica y $\lim_{n \to \infty} \lambda_n = \lambda$, entonces $\lim_{n \to \infty} \lambda_n x = \lambda x$ para cualquier $x \in X$.

Como espacios lineas con convergencia pueden indicarse los espacios lineales normalizados; no obstante, existen espacios lineales con convergencia en los que no

*) S. L. Sóbolev (n. 1908), matemático soviético; Schwartz L. (n. 1915), matemático francés.

**) Fréchet M. R. (1878—1973), matemático francés.

puede introducirse una norma que engendra la convergencia dada de sucesiones.

Definición 3. *Las aplicaciones del espcio lineal X en el conjunto de números reales R (o en el conjunto de números complejos C) reciben el nombre de funcionales definidas en dicho espacio o funcionales sobre dicho espacio. El valor de la funcional f en el punto x del espacio lineal X se designa con (f, x), es decir, igual que el producto escalar de los elementos f y x en el espacio lineal X provisto de un producto escalar.*

Esta designación se justifica, en particular, por el hecho de que el producto escalar (y, x) es, cuando el elemento y es fijo, una funcional definida sobre el espacio mencionado X.

Definición 4. *Sea X un espacio lineal. La funcional f definida sobre este espacio, se llama lineal (con más precisión, lineal homogénea), si para cualesquiera elementos $x \in X$, $y \in X$ y cualesquiera números λ, μ se cumple la condición*

$$(f, \lambda x + \mu y) = \lambda(f, x) + \mu(f, y).$$

Definición 5. *La funcional f, definida sobre un espacio lineal X con convergencia, se denomina continua, si para toda sucesión convergente $x_n \in X$, $\lim\limits_{n \to \infty} x_n = x$, se cumple la condición*

$$\lim_{n \to \infty} (f, x_n) = (f, x).$$

Las funcionales, como cualesquiera funciones numéricas, pueden sumarse, multiplicarse una por la otra, en particular, por un número. Por ejemplo, si f y g son las funcionales, entonces el valor de la funcional $\alpha f + \beta g$ (α y β son unos números) en el punto $x \in X$ se determina según la fórmula

$$(\alpha f + \beta g, x) = \alpha(f, x) + \beta(g, x).$$

Lema 1. *Las funcionales continuas lineales forman un espacio lineal.*

DEMOSTRACIÓN. Sean f y g las funciones lineales y sean α y β unos números. Mostremos que $\alpha f + \beta g$ es también una funcional lineal:

$$\begin{aligned}(\alpha f + \beta g, \lambda x + \mu y) &= \alpha(f, \lambda x + \mu y) + \beta(g, \lambda x + \mu y) = \\ &= \alpha[\lambda(f, x) + \mu(f, y)] + \beta[\lambda(g, x) + \mu(g, y)] = \\ &= \lambda[\alpha(f, x) + \beta(g, x)] + \mu[\alpha(f, y) + \beta(g, y)] = \\ &= \lambda(\alpha f + \beta g, x) + \mu(\alpha f + \beta g), y),\end{aligned}$$

es decir, $\alpha f + \beta g$ es una funcional lineal.

Supongamos ahora que f y g son funcionales continuas. Probemos que en este caso $\alpha f + \beta g$ es también una funcional continua. Sea $\lim\limits_{n \to \infty} x_n = x$. Entonces

$$\begin{aligned}\lim_{n \to \infty} (\alpha f + \beta g, x_n) &= \lim_{n \to \infty} [\alpha(f, x_n) + \beta(g, x_n)] = \\ &= \alpha \lim_{n \to \infty} (f, x_n) + \beta \lim_{n \to \infty} (g, x_n) = \alpha(f, x) + \beta(g, x) = (\alpha f + \beta g, x).\end{aligned}$$

De este modo, en un conjunto de funcionales continuas lineales están definidas de un modo natural las operaciones de su sumación y multiplicación por un núme-

ro. El cumplimiento, para dichas operaciones, de los axiomas del espacio lineal se comprueba sin dificultades algunas. □

En un espacio lineal de funcionales continuas lineales del espacio X el concepto de convergencia de las sucesiones se determina de la manera siguiente.

Definición 6. *Una sucesión de funcionales f_n, $n = 1, 2, \ldots$ se denomina convergente hacia la funcional f, si la sucesión de los valores de las funcionales f_n converge en todo punto $x \in X$ al valor de la funcional f en este punto, en otras palabras, si para cualquier elemento $x \in X$ la sucesión numérica $\{(f_n, x)\}$ converge hacia el número (f, x).*

De este modo, la afirmación $\lim\limits_{n \to \infty} f_n = f$ es equivalente a la siguiente:

$$\lim_{n \to \infty} (f_n, x) = (f, x) \text{ para todo } x \in X.$$

Definida de este modo la convergencia de las funcionales, las operaciones de su sumación y multiplicación por un número son continuas (esto se deduce inmediatamente de la linealidad de funcionales y de la propiedad de los límites de las sucesiones numéricas), y, por consiguiente, si introducimos el concepto de convergencia de las funcionales conforme a la definición 6, resultará lícita la siguiente afirmación que se enunciará en forma de un lema.

Lema 2. *Las funcionales continuas lineales, definidas sobre un espacio con convergencia, forman también un espacio lineal con convergencia.*

Definición 7. *Un espacio lineal con convergencia cuyos elementos son las funcionales continuas lineales, definidas sobre el espacio X, se llama espacio conjugado de X.*

Sean X e Y los espacios lineales con convergencia, con la particularidad de que todo elemento del espacio X es un elemento del espacio Y y supongamos que cualquier sucesión $x_n \in X$, $n = 1, 2, \ldots$, convergente en X hacia el elemento x, converge hacia x también en el espacio Y. En este caso se escribirá

$$X \hookrightarrow Y.$$

Definición 8. *Se dice que una funcional continua lineal f, definida sobre el espacio $X \hookrightarrow Y$, es prolongable al espacio Y en una funcional continua lineal, si existe tal funcional continua lineal F, definida sobre el espacio Y, que $(F, x) = (f, x)$ para todo $x \in X$ (es decir, $F = f$ en Y). En este caso la funcional F lleva el nombre de prolongación de la funcional f.*

Ejercicio 1. Sean X e Y los espacios lineales con convergencia. Demuéstrese que si $X \hookrightarrow Y$ y el conjunto X es denso en el espacio Y (es decir, cada elemento del espacio Y es el límite en este espacio para la sucesión de elementos de X), entonces cualquier funcional continua lineal del espacio X, si se prolonga en una funcional continua lineal del espacio Y, es prolongable de un modo único.

Igual que para las aplicaciones de cualesquiera espacios lineales, para los espacios con convergencia tiene sentido el concepto de aplicación lineal (operator lineal) de un espacio con convergencia en otro espacio de este mismo género (véase la definición 17 en el p. 57.2). Introduzcamos, además, el concepto de aplicación continua de un espacio lineal con convergencia en otro espacio con convergencia.

Definición 9. *Sean X_1 y X_2 dos espacios lineales con convergencia. La aplicación Φ del espacio X_1 en X_2 (en este caso la aplicación se denomina también operador) se llama continua en el punto $x_0 \in X_1$, si, cualquiera que sea la sucesión $x_n \in X_1$, $n = 1, 2, \ldots$, convergente en el espacio X_1 hacia el punto x_0, la sucesión $\Phi(x_n) \in X_2$, $n = 1, 2, \ldots$, converge en X_2 al elemento $\Phi(x_0)$.*

En otras palabras, la aplicación Φ es continua en el punto x_0, si de $\lim\limits_{n \to \infty} x_n = x_0$ proviene que $\lim\limits_{n \to \infty} \Phi(x_n) = \Phi(x_0)$.

Lema 3. *Si la aplicación lineal Φ del espacio lineal con convergencia X_1 en el espacio lineal con convergencia X_2 es continua en el cero del espacio X_1, será continua en todo punto de X_1.*

DEMOSTRACIÓN. Sea $x_0 \in X$ y $\lim\limits_{n \to \infty} x_n = x_0$; entonces $\lim\limits_{n \to \infty} (x_n - x_0) = 0$. Por cuanto la aplicación Φ es continua en el cero,

$$\lim_{n \to \infty} \Phi(x_n - x_0) = 0.$$

Ya que la aplicación Φ es lineal, se tiene

$$\Phi(x_n - x_0) = \Phi(x_n) - \Phi(x_0)$$

y, por consiguiente,

$$\lim_{n \to \infty} [\Phi(x_n) - \Phi(x_0)] = 0, \quad \text{de donde} \quad \lim_{n \to \infty} \Phi(x_n) = \Phi(x_0).$$

De este modo, la aplicación Φ es continua en todo punto $x_0 \in X_1$. □

Definición 10. *La aplicación Φ de un espacio lineal con convergencia X_1 en otro espacio lineal con convergencia X_2 se denomina continua en X_1, si es continua en todo punto del espacio X_1.*

Para todo espacio lineal con convergencia X tienen sentido los conceptos de serie $\sum\limits_{n=1}^{\infty} u_n$, $u_n \in X$, $n = 1, 2, \ldots$, de serie convergente y de suma de ésta. Estos conceptos se introducen por analogía con el caso de los espacios lineales normalizados. Esto resulta posible, por cuanto en las definiciones correspondientes de las propiedades de una norma se emplea sólo aquella que atestigua que en todo espacio normalizado queda definido el concepto de sucesión convergente.

Los ejemplos de aplicaciones lineales y continuas de los espacios con convergencia se darán en el p. 59.6 y en el 59.7.

59.3. DEFINICIÓN DE LAS FUNCIONES GENERALIZADAS. ESPACIOS D Y D'

Definamos, ante todo, un espacio lineal de funciones D que será el concepto principal en nuestros razonamientos. Con este fin consideraremos las funciones que vienen dadas en el conjunto de todos los números reales R y que toman valores complejos.

El espacio D, que nos interesa, se compone de las funciones finitas infinitamente derivables (véase en el p. 55.2 la definición de funciones finitas). Todas las fun-

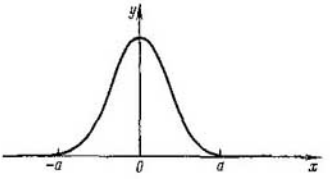


Fig. 244

ciones finitas, siendo definidas del modo natural las operaciones de su sumación y multiplicación por un número, forman un espacio lineal, y las funciones finitas infinitamente derivables forman un subespacio de dicho espacio. Introduzcamos en este subespacio el concepto de convergencia de las sucesiones.

Definición 11. *Una sucesión de las funciones finitas infinitamente derivables* φ_n, $n = 1, 2, \ldots$, *se llama convergente hacia la función finita infinitamente derivable* φ, *si*:

1) *existe un segmento* $[a, b]$, *fuera del cual todas las funciones* ψ_n, $n = 1, 2, \ldots$, *y* ψ *se anulan* *);

2) *en dicho segmento la sucesión de funciones* φ_n, $n = 1, 2, \ldots$, *y las sucesiones de todas sus derivadas* $\varphi_n^{(k)}$, $n = 1, 2, \ldots$, *convergen uniformemente hacia la función* φ *y a sus correspondientes derivadas* $\varphi^{(k)}$, $k = 1, 2, \ldots$, *respectivamente.*

Una totalidad de las funciones finitas infinitamente derivables con la operación introducida de paso límite constituye un espacio lineal con convergencia. Esto se deduce inmediatamente de las propiedades de los límites de funciones y de las propiedades de las sucesiones uniformemente convergentes.

Definición 12. *Un espacio de funciones finitas infinitamente derivables con convergencia introducida lleva el nombre de espacio D de funciones principales.*

Evidentemente, si $\varphi \in D$, toda derivada de la función φ pertenecerá también al espacio D.

Diremos además que si $\{\varphi_n\}$ converge hacia φ en D, entonces la sucesión $\{\varphi_n^{(k)}\}$ de las derivadas de cualquier orden $k = 1, 2, \ldots$ converge hacia $\varphi^{(k)}$ en D. Esta propiedad proviene directamente de la definición de convergencia en el espacio D.

Como ejemplo trivial de una función del espacio D sirve la función igual a cero en todo el eje y como ejemplo menos trivial, la función $\varphi_a(x)$ (fig. 244).

$$\varphi_a(x) = \begin{cases} e^{\frac{a^2}{a^2 - x^2}}, & \text{si} \quad |x| < a, \\ 0, & \text{si} \quad |x| \geqslant a. \end{cases} \tag{59.13}$$

*) El segmento $[a, b]$ contiene los portadores de todas las funciones φ, φ_n, $n = 1, 2, \ldots$.

Ejercicios. 2. Demuéstrese que la función (59.13) es infinitamente derivable en todo el eje numérico (compárese con (37.25)).

3. Demuéstrese que con miras de lograr que para la función $\varphi \in D$ exista una función $\psi \in D$ tal que sea $\varphi = \psi'$, es necesario y suficiente que $\int_{-\infty}^{+\infty} \varphi(t)\, dt = 0$.

Definición 13. *Toda funcional continua lineal f, definida sobre D, recibe el nombre de función generalizada.*

Definición 14. *Una funcional f, definida en todo el eje real, se denomina localmente integrable, si es absolutamente integrable en cualquier segmento finito.*

Si f es una función localmente integrable y $\varphi \in D$, el producto f_φ es absolutamente integrable en todo el eje. En efecto, sea supp $\varphi \subset [a, b]$ (véase en el p. 55.2 la definición del portador supp φ de la función φ); la función φ es, obviamente, acotada: $|\varphi(x)| \leqslant C$, $-\infty < x < +\infty$, por lo cual

$$\int_{-\infty}^{+\infty} |f(x)\varphi(x)|\, dx = \int_a^b |f(x)\varphi(x)|\, dx \leqslant C \int_a^b |f(x)|\, dx.$$

Definamos, para la función localmente integrable f, la funcional (f, φ) sobre D mediante la igualdad

$$(f, \varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)\, dx. \tag{59.14}$$

Esta funcional es lineal y continua. Su linealidad es obvia; demostremos su continuidad. Sea $\lim_{n\to\infty} \varphi_n = \varphi$ en D. En este caso existe un segmento $[a, b]$ tal que para todo $n = 1, 2, \ldots$, tienen lugar las inclusiones supp $\varphi_n \subset [a, b]$ y supp $\varphi \subset [a, b]$; por eso

$$|(f, \varphi) - (f, \varphi_n)| \leqslant \int_{-\infty}^{+\infty} |f(x)|\,|\varphi(x) - \varphi_n(x)|\, dx =$$

$$= \int_a^b |f(x)|\,|\varphi(x) - \varphi_n(x)|\, dx \leqslant \sup_{[a,b]} |\varphi(x) - \varphi_n(x)| \int_a^b |f(x)|\, dx \to 0$$

para $n \to \infty$. De este modo, a toda función localmente integrable $f(x)$ le corresponde una función generalizada (f, φ) [*]; en este sentido toda función localmente integrable puede considerarse como una función generalizada.

Una función constante, es decir, tal función generalizada f que $(f, \varphi) = c \int_{-\infty}^{+\infty} \varphi(x)\, dx$, c es una constante, $\varphi \in D$ (en particular, la función nula), se genera por una función localmente integrable $f(x) = c$, $-\infty < x < +\infty$.

---

[*] En este caso suele decirse también que la función generalizada (f, φ) se genera por la función f.

Recordemos que por funcional lineal se entiende siempre una funcional lineal homogénea. Por esta razón, la funcional f, que es igual a una misma constante c, en todos los elementos de cierto espacio lineal X, aunque es una función lineal, no es, sin embargo, homogénea: si $x \in X$, $y \in X$, λ y μ son unos números, entonces para la funcional mencionada tendremos $f(x) = f(y) = f(\lambda x + \mu y) = c$. Si fuera lineal homogénea, debería verificarse $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = (\lambda + \mu)c$. Por cuanto para $\lambda + \mu \neq 1$ se tiene $f(\lambda x + \mu y) \neq \lambda f(x) + \mu f(y)$, entonces la funcional f no es linealmente homogénea. Por consiguiente, una funcional, igual a una constante, no pertenece a la clase de funcionales que se considera.

Ejercicio 4. Demuéstrese que dos funciones continuas en un eje numérico son distintas cuando, y sólo cuando, son distintas las funciones generalizadas engendradas por las primeras.

A veces, las funciones generalizadas se designan mediante el símbolo $f(x)$. Esta designación es simbólica pura; no significa ni mucho menos el valor de la función generalizada en el punto $x \in R$, sino sólo refleja el hecho de que las funciones generalizadas son, en el sentido señalado más arriba, una generalización de las funciones (localmente integrables) ordinarias; no se presupone ningún valor de la función generalizada en el punto x.

Para denotar el valor de la función generalizada f en el punto $\varphi = \varphi(x)$ del espacio D, a la par con la notación (f, φ) se usa también la inscripción

$$\int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx. \tag{59.15}$$

De este modo, por definición,

$$\int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx = (f, \varphi).$$

Esta igualdad constituye la definición del símbolo (59.15), el cual se lee formalmente como "integral del producto de f por φ". Dicha notación refleja el hecho de que las funciones generalizadas son una generalización de las funcionales (59.14), donde f es una función localmente integrable.

Ejercicio 5. Demuéstrese que la funcional $v.p. \int_{-\infty}^{+\infty} \frac{\varphi(x)}{x}\,dx$, $\varphi \in D$, es una función generalizada $\left(\textit{se denota habitualmente } \mathscr{P}\frac{1}{x}\right)$.

A título de otro ejemplo de una función generalizada consideraremos una funcional designada por $\delta = \delta(x)$ y llamada función delta (véase 59.1).

Definición 15. *Una funcional definida por la fórmula*

$$(\delta, \varphi) = \varphi(0), \quad \varphi \in D,$$

se denomina función delta.

La linealidad y continuidad de dicha funcional se comprueban fácilmente. No puede ser representada en la forma (59.14), cualquiera que sea la función localmen-

te integrable f. Efectivamente, si se encontrara una función localmente integrable f tal que

$$(\delta, \varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx, \quad \varphi \in D,$$

entonces para esta función f y para la función φ, dada mediante la fórmula (59.13), tendríamos

$$\int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx = \varphi(0) = \frac{1}{e}. \tag{59.16}$$

Mas, por ser la función f absolutamente integrable,

$$\lim_{a \to 0} \int_{-a}^{a} |f(x)|\,dx = 0$$

(¿por qué?).

Luego, al observar que $e^{-\frac{a^2}{a^2-x^2}} < \frac{1}{e}$, $-a \leqslant x \leqslant a$, obtenemos

$$\left| \int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx \right| = \left| \int_{-a}^{a} f(x) e^{-\frac{a^2}{a^2-x^2}}\,dx \right| \leqslant \frac{1}{e} \int_{-a}^{a} |f(x)|\,dx,$$

razón por la cual el primer miembro de la igualdad (59.16) tiende a cero cuando $a \to 0$. mientras que el segundo miembro no tiende a cero. La contradicción obtenida demuestra precisamente nuestra afirmación. De este modo, la reserva de funciones generalizadas en el sentido indicado es mayor que el de funciones ordinarias.

Definición 16. *Una funcional que a toda función $\varphi \in D$ le pone en correspondencia un número $\varphi(x_0)$, donde x_0 es fijo, también se llama función delta y se denota mediante $\delta(x - x_0)$.*

Recurriendo a la notación (59.15), podemos escribir

$$\int_{-\infty}^{+\infty} \delta(x - x_0)\varphi(x)\,dx = \varphi(x_0), \quad \varphi \in D.$$

Definición 17. *Una totalidad de funciones generalizadas, al igual que toda totalidad de funcionales, definidas sobre un espacio lineal con convergencia (véase el p. 59.2), forma un espacio lineal con convergencia conjugado de D. Se llama espacio de funciones generalizadas y se denota con D'.*

Así pues, la convergencia de una sucesión de funciones generalizadas f_n, $n = 1, 2, \ldots$, hacia la función generalizada f significa que

$$\lim_{n \to \infty} (f_n, \varphi) = (f, \varphi)$$

para cualquier función $\varphi \in D$.

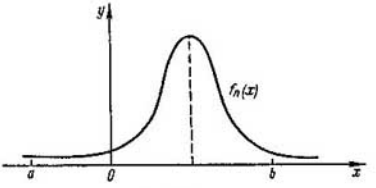


Fig. 245

Problema 40. Sea $f_n \in D'$, $n = 1, 2, \ldots$, y supongamos que para cualquier función $\varphi \in D$ existe un límite de la sucesión numérica (f_n, φ). Pongamos $F(\varphi) = \lim_{n \to \infty} (f_n, \varphi)$. Demuéstrese que $F(\varphi)$ es una función generalizada.

En el p. 59.1 se han considerado las funciones $\delta_\varepsilon(x)$ que son, con evidencia, localmente integrables. Hemos visto que poseen la propiedad de que para cualquier función continua en todo el eje φ, y por lo tanto, para cualquier función $\varphi \in D$ se tiene

$$\lim_{\varepsilon \to +0} (\delta_\varepsilon, \varphi) = \lim_{\varepsilon \to 0} \int_{-\infty}^{+\infty} \delta_\varepsilon(x)\varphi(x)\,dx = \varphi(0) = (\delta, \varphi).$$

Desde el punto de vista de las funciones generalizadas esto significa que en D'.

$$\lim_{\varepsilon \to +0} \delta_\varepsilon = \delta^{*)}.$$

De este modo, la función delta en el espacio D' es un límite de una sucesión de funciones generalizadas engendradas por las funciones localmente integrables.

Ejercicios. 6. Supongamos que la sucesión de funciones absolutamente integrables $f_n(x)$, $n = 1, 2, \ldots$, es tal que

a) cualquiera que sea el número $M > 0$, para $|a| < M$, $|b| < M$, las magnitudes

$$\left| \int_a^b f_n(x)\,dx \right|, \quad n = 1, 2, \ldots$$

están acotadas por una constante que no depende de a, b, n (sólo depende de M);

b) cualesquiera que sean a y b fijos y distintos de cero,

$$\lim_{n \to \infty} \int_a^b f_n(x)\,dx = \begin{cases} 0 & \text{para } a < b < 0 \text{ y } 0 < a < b, \\ 1 & \text{para } a < 0 < b. \end{cases}$$

Las sucesiones de este género se denominan *sucesiones en delta* (fig. 245).

---

*) Al igual que en el caso de las funciones ordinarias, el símbolo $\varepsilon \to +0$ significa que la relación límite mencionada tiene lugar para cualquier sucesión $\varepsilon_n > 0$, $n = 1, 2, \ldots$, que tiende a cero.

Demuéstrese que para cualquier función continua φ y cualquier sucesión en delta $f_n(x)$, $n = 1, 2, \ldots$,

$$\lim_{n \to \infty} \int_{-\infty}^{+\infty} f_n(x)\varphi(x)\,dx = \varphi(0);$$

en otras palabras, $\lim_{n \to \infty} (f_n, \varphi) = (\delta, \varphi)$

7. Sea $f_t(x) = \frac{1}{\sqrt{\pi t}} e^{-\frac{x^2}{t}}$. Demuéstrese que en el espacio D' se verifica la igualdad $\lim_{t \to +0} f_t(x) = \delta(x)$.

8. Demuéstrese que en el espacio D' existe el límite $\lim_{y \to 0} \frac{1}{x \pm iy}$ $\left(\text{se denota con } \frac{1}{x \pm i0}\right)$ y que son válidas las fórmulas

$$\frac{1}{x \pm i0} = \mp i\pi\delta(x) + \mathcal{P}\frac{1}{x}.$$

(estas fórmulas llevan el nombre de Sojotski *)).

Problema 41. Demuéstrese que toda función generalizada singular es el límite de funciones regulares (en este sentido un espacio de las funciones generalizadas es el "complemento" del espacio de funciones ordinarias).

Como ya hemos visto, el concepto de función generalizada no se reduce al de función de un punto, por lo cual, en general, no hay sentido en hablar del valor de una función generalizada en el punto dado, en particular, de la reducción de esta función a cero en dicho punto. Sin embargo, se puede introducir un concepto natural de reducción a cero de una función generalizada en un intervalo.

Definición 18. *Diremos que una función generalizada f se reduce a cero en el intervalo (a, b), si $(f, \varphi) = 0$ para cualquier $\varphi \in D$, que dispone de un portador contenido dentro del intervalo (a, b).*

Ejercicio 9. Demuéstrese que con el fin de lograr que una función continua se redfizca a cero en todo punto de un intervalo, es necesario y suficiente que se reduzca a cero en dicho intervalo como una función generalizada.

Definición 19. *Las funciones generalizadas f y g se llaman iguales en el intervalo (a, b), si $f - g = 0$ en (a, b).*

59.4. DERIVACIÓN DE LAS FUNCIONES GENERALIZADAS

Determinemos ahora la derivada de una función generalizada. Aclaremos, ante todo, qué representa en sí la derivada de una función ordinaria f, continuamente derivable en todo el eje numérico, cuando dicha derivada se considera como la funcional (f', φ) en D. Esto tiene sentido, por cuanto la derivada f', siendo continua en todo el eje numérico, es una función localmente integrable.

Integrando por partes, por ser finita la función $\varphi \in D$, obtendremos

*) Yu. V. Sojotski (1842—1929), matemático ruso.

$$(f', \varphi) = \int_{-\infty}^{+\infty} f'(x)\varphi(x)\,dx = -\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,dx = -(f, \varphi'), \qquad (59.17)$$

donde, como se sabe, $\varphi' \in D$. De este modo, la derivada f' es una funcional sobre D cuyos valores se expresan en términos de los valores de la función f, considerada como una funcional, por medio de la fórmula (59.17). Esto hace natural la siguiente definición.

Definición 20. *Se llama derivada de la función generalizada f una funcional sobre D, denotada con f' y determianda por la igualdad*

$$(f', \varphi) = -(f, \varphi'), \quad \varphi \in D. \qquad (59.18)$$

En otras palabras, el valor de la funcional f' en cualquier punto φ del espacio D es igual al valor de la funcional f en el punto $\varphi' \in D$ tomada con el signo opuesto.

Así pues, cualquier función generalizada tiene una derivada. De aquí proviene que toda función localmente derivable también tiene derivada en el sentido de la definición 20.

De la fórmula (59.17) se deduce que la derivada en el sentido habitual de una función continuamente derivable en todo el eje numérico, considerada como funcional sobre D, coincide con la derivada de dicha función en el sentido de las funciones generalizadas.

La operación de cálculo de la derivada de una función generalizada la llaman, por analogía con el caso de las funciones ordinarias, derivación.

Lema 4. *La funcional f' es lineal continua y, por lo tanto, constituye una función generalizada.*

DEMOSTRACIÓN. Comprobemos la linealidad:

$$(f', \lambda\varphi + \mu\psi) = -(f, (\lambda\varphi + \mu\psi')) = -(f, \lambda\varphi' + \mu\psi') =$$
$$= -\lambda(f, \varphi') - \mu(f, \psi') = \lambda(f', \varphi) + \mu(f', \psi), \quad \varphi \in D, \quad \psi \in D.$$

Para comprobar la continuidad de la funcional f, recordemos que si $\varphi \in D$, $\varphi_k \in D$, $k = 1, 2, \ldots$, y $\lim_{k\to\infty} \varphi_k = \varphi$ en D, entonces, en virtud de la definición de convergencia en el espacio D, también $\lim_{k\to\infty} \varphi_k' = \varphi'$ en D; por ello, si $\varphi_k \to \varphi$ en D, entonces $\lim_{k\to\infty} (f', \varphi_k) = -\lim_{k\to\infty} (f, \varphi_k') = -(f, \varphi') = (f', \varphi)$.

De este modo, si $f \in D'$, entonces f' existe siempre y $f' \in D'$. □

Las derivadas de órdenes superiores de una función generalizada se determinan sucesivamente, al igual que en el caso de las funciones ordinarias:

$$f'' = (f')', \quad f''' = (f'')', \ldots,$$

en general

$$f^{(k)} = (f^{(k-1)})' \quad k = 1, 2, \ldots, f^{(0)} = f.$$

Por inducción se comprueba con facilidad que

$$(f^{(k)}, \varphi) = (-1)^k (f, \varphi^{(k)}), \quad \varphi \in D, \quad k = 0, 1, \ldots.$$

Conforme a esta definición, las funciones generalizadas tienen derivadas de cualesquiera órdenes o, como se dice a veces, son infinitamente derivables.

Ejemplos. 1. Sea

$$\theta(x) = \begin{cases} 1, & \text{si } x \geqslant 0, \\ 0, & \text{si } x < 0. \end{cases}$$

La función $\theta(x)$ lleva el nombre *de Heaviside* (véase (59.10)) o de función unidad. Es localmente integrable y por esta razón puede considerarse como una función generalizada. Hallemos su derivada. De acuerdo con la definición (59.18),

$$(\theta', \varphi) = -(\theta, \varphi') = -\int_{-\infty}^{+\infty} \theta(x)\varphi'(x)\,dx = -\int_{0}^{+\infty} \varphi'(x)\,dx =$$

$$= \varphi(0) = (\delta, \varphi), \quad \varphi \in D,$$

es decir, $\theta' = \delta$ (compárese con el p. 59.1).

2. Calculemos, a título de otro ejemplo, las derivadas de la función delta

$$(\delta^{(k)}, \varphi) = (-1)^k(\delta, \varphi^{(k)}) = (-1)^k\varphi^{(k)}(0).$$

Ejercicios. 10. Sean f y g las funciones generalizadas y sean λ y μ unos números. Demuéstrese que

$$(\lambda f + \mu g)' = \lambda f' + \mu g'.$$

11. Demuéstrese que $\left(\frac{d}{dx} + \lambda\right)\theta(x)e^{-\lambda x} = \delta(x)$.

12. Demuéstrese que $\left(\frac{d^2}{dx^2} + \omega^2\right)\frac{\theta(x)\operatorname{sen}\omega x}{\omega} = \delta(x)$.

13. Si $\delta_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon} & \text{para } |x| < \frac{\varepsilon}{2}; \\ 0 & \text{para } |x| \geqslant \frac{\varepsilon}{2}; \end{cases}$, entonces en el espacio de las funciones generalizadas

$$\lim_{\varepsilon \to 0} \delta_\varepsilon(x) = \delta(x) \quad \text{y} \quad \delta'_\varepsilon(x) = \frac{\delta\left(x + \frac{\varepsilon}{2}\right) - \delta\left(x - \frac{\varepsilon}{2}\right)}{\varepsilon}.$$

14. Sea $f(x) = \begin{cases} f_1(x) \text{ para } x < x_0, \\ f_2(x) \text{ para } x > x_0 \end{cases}$ donde las funciones $f_1(x)$ y $f_2(x)$ son continuamente derivables y existen los límites $f'(x_0 \pm 0)$. Hállese la derivada $f'(x)$ en el espacio D

15. Sea $f(x)$ una función continuamente derivable en todo el eje numérico. Hállese la derivada $(\theta f)'$ en el espacio D'.

16. Demuéstrese que en el espacio D' es válida la fórmula $\mathscr{D}\frac{1}{x} = (\ln|x|)'$ (véase el ejercicio 5).

17. Demuéstrese que en el espacio D' se verifica la igualdad

$$\sum_{n=1}^{\infty} \cos nx = -\frac{1}{2} + \pi \sum_{k=-\infty}^{+\infty} \delta(x - 2k\pi).$$

Indicación: hágase uso de la fórmula (véase el ejemplo 3 en el p. 55.4)

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n} = \frac{\pi - x}{2}, \quad 0 < x < 2\pi.$$

Lema 5. *Sea* $f_n \in D'$, $f \in D'$ *y*

$$\lim_{n \to \infty} f_n = f; \tag{59.19}$$

en este caso también

$$\lim_{n \to \infty} f_n' = f', \tag{59.20}$$

es decir, para toda sucesión de funciones generalizadas convergente en D' *la derivada de la función límite es igual al límite de la sucesión de las derivadas.*

DEMOSTRACIÓN. Para cualquier función $\varphi \in D$

$$(f', \varphi) - (f_n', \varphi) = -[(f, \varphi') - (f_n, \varphi')] \to 0 \text{ para } n \to \infty,$$

pues $\varphi' \in D$. □

Pueden considerarse también las series de las funciones generalizadas

$$\sum_{n=1}^{\infty} u_n, \tag{59.21}$$

donde $u_n \in D'$, $n = 1, 2, \ldots$. La suma

$$s_n = \sum_{k=1}^{n} u_k$$

se llama *suma parcial de n-ésimo orden* ($n = 1, 2, \ldots$) de la serie (59.21). La serie (59.21) se llama *convergente*, si en D' existe el límite

$$\lim_{n \to \infty} s_n = s.$$

La función generalizada s lleva el nombre de suma de la serie (59.21); se escribe en este caso

$$s = \sum_{n=1}^{\infty} u_n.$$

Lema 6. *Una serie convergente de las funciones generalizadas puede derivarse término por término tantas veces como se quiera:*

$$s^{(k)} = \sum_{n=1}^{\infty} u_n^{(k)}, \; k = 1, 2, \ldots.$$

Esto se deduce del lema 5.

59.5. ESPACIO DE LAS FUNCIONES PRINCIPALES S Y ESPACIO DE LAS FUNCIONES GENERALIZADAS S'.

Designemos con S el conjunto de todas las funciones de valores complejos infinitamente derivables en todo el eje numérico, las cuales, al igual que todas las derivadas suyas, tienden a cero, para $x \to \infty$, más rápido que cualquier potencia de $\frac{1}{x}$. En otras palabras, el conjunto S se compone de aquellas, y sólo aquellas, funciones infinitamente derivables φ, para las cuales se cumple la condición

$$\lim_{x \to \infty} x^n \varphi^{(m)}(x) = 0, \tag{59.22}$$

cualesquiera que sean n y m enteros y no negativos. La condición de pertenencia de la función φ al conjunto S puede formularse de una manera algo diferente: una función infinitamente derivable φ pertenece a S cuando, y sólo cuando, para cualesquiera n y m enteros y no negativos se tiene

$$\sup_{-\infty < x < +\infty} |x^n \varphi^{(m)}(x)| = c_{n,m} < \infty. \tag{59.23}$$

En efecto, si esto es así, entonces, al sustituir en (59.23) n por $n + 1$, obtendremos $|x^{n+1}\varphi^{(m)}(x)| \leqslant c_{n+1,m}$, por lo cual

$$|x^n \varphi^{(m)}(x)| \leqslant \frac{c_{n+1,m}}{|x|},$$

de donde proviene (59.22). Viceversa, de (59.22) y del hecho de que $x^n\varphi^{(m)}(x)$ es acotada en cualquier segmento, se deduce (59.23).

Es evidente que el conjunto S es un espacio lineal. En tal caso si $\varphi \in S$, entonces cualquier derivada de la función φ también pertenece al espacio S.

Definición 21. *Una sucesión de funciones $\varphi_k(x) \in S$, $k = 1, 2, \ldots$, se llama convergente en S hacia la función $\varphi(x) \in S$, si para cualesquiera n y m enteros y no negativos toda sucesión $x^n\varphi_k^{(m)}(x)$, $k = 1, 2, \ldots$, converge uniformemente en todo el eje hacia la función $x^n\varphi^{(m)}(x)$.*

Es evidente que $\lim_{k \to \infty} \varphi_k = \varphi$ en S cuando, y sólo cuando, para cualesquiera n y m enteros y no negativos se verifica

$$\lim_{k \to \infty} \sup_{-\infty < x < +\infty} |x^n[\varphi_k^{(m)}(x) - \varphi^{(m)}(x)]| = 0. \tag{59.24}$$

Ha de notarse que si $\varphi_k \to \varphi$ en S, entonces para las derivadas de cualquier orden $\varphi_k^{(m)} \to \varphi^{(m)}$ en S, $m = 1, 2, \ldots$. El espacio lineal S con la operación introducida de paso límite será un espacio lineal con convergencia.

Evidentemente, $D \hookrightarrow S$, en particular, la sucesión de funciones $\varphi_k \in D$, $k = 1, 2, \ldots$, convergente en D hacia la función φ, converge a la función φ también en S. Además, $D \neq S$, pues $e^{-x^2} \in S$, pero $e^{-x^2} \notin D$.

Problema 42. Demuéstrese que el espacio D es denso en S, es decir, cualquier función $\varphi \in S$ es un límite en S para cierta sucesión de funciones $\varphi_k \in D$, $k = 1, 2, \ldots$.

Definición 22. *Una funcional continua lineal, definida sobre el espacio S, se denomina función generalizada de crecimiento lento. Un conjunto de todas las fun-*

cionales de esta índole se llama espacio de funciones generalizadas de crecimiento lento y se denota por S′.

Toda funcional $f \in S'$, considerada sólo en el conjunto D, es una función generalizada, por consiguiente, un elemento del conjunto S' puede ser interpretado como prolongación de cierta funcional continua lineal desde el conjunto D a S (véase el p. 59.2). Por ejemplo, la funcional δ, definida en el p. 59.3 sobre el espacio D mediante la fórmula $(\delta, \varphi) = \varphi(0)$, $\varphi \in D$, puede ser prolongada con ayuda de la misma fórmula al espacio S.

Se puede mostrar que no toda función generalizada de D' puede ser prolongada a S y en este sentido podemos decir que S' constituye una parte escricta de D'.

Ejercicio 18. Demuéstrese que una función generalizada, engendrada por la función localmente integrable e^x, no puede ser prolongada al elemento del espacio S

Toda función localmente integrable $f(x)$, para la cual en cierto entorno de ∞ se verifica la estimación

$$|f(x)| \leqslant A|x|^k \tag{59.25}$$

(A y k son unas constantes no negativas) *), en particular, cualquier polinomio engendra una funcional del espacio D prolongable a la funcional continua lineal sobre S. Ésta se determina según la fórmula

$$(f, \varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx, \quad \varphi \in S. \tag{59.26}$$

Efectivamente, de las condiciones (59.22) y (59.25) proviene que $f(x)\varphi(x) \to 0$, para $x \to \infty$, más rápido que cualquier potencia de $\frac{1}{x}$, y por lo tanto, la integral (59.26) existe.

Demos a conocer, además, que toda función $f(x)$, absolutamente integrable en todo el eje numérico, genera también, según la fórmula (59.26), una funcional continua lineal sobre S. En efecto, por cuanto toda función $\varphi \in S$ está acotada, la existencia de la integral (59.26) en este caso se infiere de la desigualdad

$$\int_{-\infty}^{+\infty} |f(x)\varphi(x)|\,dx \leqslant \sup_{-\infty<x<+\infty} |\varphi(x)| \int_{-\infty}^{\infty} |f(x)|\,dx.$$

Ejercicios. 19. Demuéstrese que la funcional (59.26) es lineal y continua sobre el espacio S (tanto en el caso en que la función f és de crecimiento lento en el infinito, como cuando sea absolutamente integrable en todo el eje numérico).

20. Demuéstrese que la función generalizada $\frac{1}{x + i0} \in D'$ (véase el ejercicio 8) es prolongable al elemento de espacio S

*) Tales funciones se llaman funciones de crecimiento lento, de donde proviene el término "funciones generalizadas de crecimiento lento"

El conjunto S' forma un espacio lineal con convergencia, conjugado de S (véase el p. 59.2).

Por cuanto para cualquier función $\varphi \in S$ tendremos $\varphi' \in S$, entonces para las funciones generalizadas del espacio S', al igual que para las funciones generalizadas de D', puede hallarse la derivada f' según la fórmula

$$(f', \varphi) = -(f, \varphi'), \quad \varphi \in S.$$

De este modo, para cualquier función generalizada $f \in S'$ la derivada f' siempre existe y $f' \in S'$. Además, las derivadas de la función generalizada f en el elemento $\varphi \in D$, consideradas como derivadas en los espacios D' y S', respectivamente, coinciden. Igual que en el caso del espacio D', en el espacio S' la derivada del límite siempre existe y es igual al límite de las derivadas.

59.6. TRANSFORMACIÓN DE FOURIER EN EL ESPACIO S

Toda función $\varphi \in S$ es absolutamente integrable. Más aún, si $\varphi \in S$, entonces, cualquiera que sea $k = 1, 2, \ldots$, la función $x^k\varphi(x)$ es también absolutamente integrable en todo el eje numérico. En efecto, por cuanto para la función $\varphi \in S$ se cumple la condición (59.23), se tiene

$$|x^k\varphi(x)| \leqslant c_{k,0},$$
$$x^2|x^k\varphi(x)| = |x^{k+2}\varphi(x)| \leqslant c_{k+2,0}$$

y, por esta razón,

$$|x^k\varphi(x)| \leqslant \frac{c_{k,0} + c_{k+2,0}}{1 + x^2}. \tag{59.27}$$

En el segundo miembro de esta desigualdad figura una función absolutamente integrable en todo el eje numérico, por consiguiente, de acuerdo con el criterio de comparación para las integrales impropias, la función $x^k\varphi(x)$ es también absolutamente integrable para todo $k = 0, 1, 2, \ldots$. De aquí se desprende que para las funciones $\varphi \in S$ existe la transformación clásica de Fourier

$$\hat{\varphi} = F[\varphi] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x)e^{-ixy}\,dx, \varphi \in S \tag{59.28}$$

como también la transformación de Fourier inversa

$$F^{-1}[\varphi] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(y)e^{ixy}\,dy, \quad \varphi \in S.$$

El carácter clásico de la transformación de Fourier se entiende aquí en el sentido de que las integrales escritas son ordinarias y absolutamente convergentes, y no son integrales en el sentido del valor principal (véase el p. 56.3). Sobre S son válidas las fórmulas de inversión para la transformación directa e inversa de Fourier (véase el p. 56.5):

$$F[F^{-1}[\varphi]] = \varphi, \quad F^{-1}[F[\varphi]] = \varphi, \quad \varphi \in S. \tag{59.29}$$

Observemos que, por ejemplo, la segunda de estas fórmulas en la forma integral tiene por expresión

$$F^{-1}[\hat{\varphi}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{\varphi}(y) e^{ixy}\, dy = \varphi(x).$$

Teorema 1. *La transformación de Fourier y la transformación de Fourier inversa aplican el espacio S sobre sí de manera biunívoca, lineal y continua.*

DEMOSTRACIÓN. Señalemos que si $\varphi \in S$, entonces también $\hat{\varphi} \in S$.

Ante todo, de lo que para toda función $\varphi \in S$ la función $x^k \varphi(x)$ con $k = 0, 1, 2, \ldots$ es, según se ha probado anteriormente, absolutamente integrable en todo el eje numérico, proviene, de acuerdo con el teorema 4 del p. 56.10, que la transformación de Fourier $\hat{\varphi} = F[\varphi]$ de la función φ existe y represneta en sí una función infinitamente derivable.

Estimemos ahora la función $|y^n \hat{\varphi}^{(m)}(y)|$, donde n y m son los números enteros no negativos. Aplicando las fórmulas para la derivada de la transformación de Fourier (véase el p. 56.10) y para la transformación de Fourier de la derivada (véase el p. 56.8), obtendremos

$$|y^n \hat{\varphi}^{(m)}(y)| = |y^n F^{(m)}[\varphi]| = |y^n F[x^m \varphi]| =$$

$$= |F[(x^m \varphi)^{(n)}]| = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x^m \varphi(x))^{(n)} e^{-ixy}\, dx \right| \leqslant$$

$$\leqslant \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |(x^m \varphi(x))^{(n)}|\, dx.$$

Observemos que la expresión $[x^m \varphi(x)]^{(n)}$ representa, en virtud de las reglas de derivación, una combinación lineal de las expresiones del tipo $x^p \varphi^{(q)}(x)$, donde p y q son enteros no negativos y, de acuerdo con lo notado anteriormente, $\varphi^{(q)} \in S$. Por ello (véase (59.27)), las funciones $(1 + x^2) x^p \varphi^{(q)}$ están acotadas en todo el eje numérico, por lo cual está acotada también la función $(1 + x^2)\, [x^m \varphi(x)]^{(n)}$, es decir,

$$\sup_{-\infty < x < +\infty} (1 + x^2) |[x^m \varphi(x)]^{(n)}| < +\infty.$$

Ahora, dividamos y multipliquemos la expresión subintegral obtenida por $1 + x^2$, entonces, teniendo presente que $\int_{-\infty}^{+\infty} \frac{dx}{1 + x^2} = \pi$, llegamos a que

$$|y^n \hat{\varphi}^{(m)}(y)| \leqslant \frac{1}{\sqrt{2\pi}} \sup_x (1 + x^2) |(x^m \varphi(x))^{(n)}| \int_{-\infty}^{+\infty} \frac{dx}{1 + x^2} =$$

$$= \sqrt{\frac{\pi}{2}} \sup_x (1 + x^2) |(x^m \varphi(x))^{(n)}|. \tag{59.30}$$

Ya que a la derecha figura una magnitud finita, entonces $\hat{\varphi} \in S$.

Así pues, la transformación de Fourier aplica S en S y esta aplicación es biunívoca (véase el lema 3 en el p. 56.5).

Análogamente se demuestra también que la transformación de Fourier inversa F^{-1} aplica S en S y, además, biunívocamente. Es fácil convencerse de que estas aplicaciones se realizan en realidad sobre el espacio S, es decir, son las biyecciones. Esto se deduce directamente de las fórmulas de reciprocidad (59.29) para las transformaciones directa e inversa de Fourier *).

Efectivamente, probemos que $F(S)$ coincide con todo el espacio S. Sea $\psi \in S$ y pongamos $\varphi = F^{-1}[\psi]$.

En este caso

$$F[\varphi] = F[F^{-1}[\psi]] = \psi.$$

De un modo semejante se demuestra también que

$$F^{-1}(S) = S.$$

La linealidad de la transformación de Fourier se ha demostrado antes (véase el lema 2 en el p. 56.5).

Demostremos ahora la continuidad de la aplicación F.

Probemos al principio su continuidad en el cero. Sea $\lim_{k \to \infty} \varphi_k = 0$ en S. Entonces, de (59.30) se deduce que

$$|y^n \hat{\varphi}_k^{(m)}(y)| \leqslant \sqrt{\frac{\pi}{2}} \sup_x (1 + x^2) |x^m \varphi_k(x))^{(n)}|, \quad k = 1, 2, \ldots .$$

Pero de (59.24) (para $\varphi(x) = 0$) tenemos

$$\lim_{k \to \infty} \sup_x (1 + x^2) |(x^m \varphi_k(x))^{(n)}| = 0;$$

por lo cual

$$\lim_{k \to \infty} \sup_y |y^n \hat{\varphi}_k^{(m)}(y)| = 0, \text{ es decir, } \lim_{k \to \infty} \hat{\varphi}_k = 0 \text{ en } S.$$

Por cuanto la transformación de Fourier es una aplicación lineal del espacio lineal S en sí mismo, continua en el cero, será continua también en todos los puntos de este espacio (véase el lema 3 en el p 59.2).

De este modo, la transformación de Fourier F aplica continuamente S sobre S.

Del modo enteramente análogo se demuestra la continuidad de la transformación de Fourier inversa F^{-1}. □

---

*) Observemos en adición que de lo que $F(S) = F^{-1}(S) = S$ se deduce que en las fórmulas (59.29) todas las integrales existen en el sentido habitual, y no sólo en el sentido del valor principal (compárese con el p. 56.5)

59.7. TRANSFORMACIÓN DE FOURIER DE LAS FUNCIONES GENERALIZADAS

Demostremos previamente una igualdad integral. Supongamos que la función f es continua y absolutamente integrable en todo el eje numérico y sea $\varphi \in S$, entonces

$$\int_{-\infty}^{+\infty} \varphi(x)\,dx \int_{-\infty}^{+\infty} f(y)e^{-ixy}\,dy = \int_{-\infty}^{+\infty} f(y)\,dy \int_{-\infty}^{+\infty} \varphi(x)e^{-ixy}\,dx. \tag{59.31}$$

Esto se deduce del teorema 7 del p. 54.3. En efecto, la integral reiterada en el primer miembro existe, pues existe la integral

$$\int_{-\infty}^{+\infty} \left| \varphi(x) \int_{-\infty}^{+\infty} f(y)e^{-ixy}\,dy \right| dx \leqslant \int_{-\infty}^{+\infty} |\varphi(x)|dx \int_{-\infty}^{+\infty} |f(y)|\,dy.$$

Si $[a, b]$ es un segmento arbitrario, la función f, siendo continua, está acotada en $[a, b]$: $|f(x)| \leqslant M$; por esta razón

$$|f(y)\varphi(x)e^{-ixy}| \leqslant M|\varphi(x)|, \quad a \leqslant y \leqslant b.$$

De aquí, por ser convergente la integral $\int_{-\infty}^{+\infty} |\varphi(y)|\,dx$, proviene la convergencia uniforme de la integral $f(y) \int_{-\infty}^{+\infty} \varphi(x)e^{-ixy}\,dx$ en el segmento $[a, b]$.

Luego, $|\varphi(x)| \leqslant c_{0,0}$, $-\infty < x < +\infty$ (véase (59.23)); por eso $|\varphi(x)f(y)e^{-ixy}| \leqslant c_{0,0}\,|f(y)|$, y, como la integral $\int_{-\infty}^{+\infty} |f(y)|\,dy$ converge, la integral

$$\varphi(x) \int_{-\infty}^{+\infty} f(y)e^{-ixy}\,dy$$

converge uniformemente en todo el eje.

Por fin, la integral

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} |\varphi(x)f(y)e^{-ixy}|\,dy = \int_{-\infty}^{+\infty} |\varphi(x)|\,dx \int_{-\infty}^{+\infty} f(y)\,dy$$

es finita, razón por la cual en el caso que se considera están cumplidas todas las condiciones del teorema 7 (p. 54.3) y, por ende, puede permutarse el orden de integración. La igualdad (59.31) queda demostrada.

Si la función $F[f]$ engendra cierta funcional sobre S (por ejemplo, satisface la condición (59.25) o es absolutamente integrable en todo el eje numérico), entonces, al multiplicar la igualdad (59.31) por $1/\sqrt{2\pi}$, obtendremos

$$(F[f], \varphi) = (f, F[\varphi]), \quad \varphi \in S. \tag{59.32}$$

Esta fórmula la tomaremos como definición de la transformación de Fourier de funciones generalizadas pertenecientes al espacio S'

Definición 23. *Se llama transformación de Fourier de una función generalizada $f \in S'$ una funcional $F[f]$ definida por la fórmula* (59.32).

Así pues, para cualquier función generalizada f de S' está definida su transformación de Fourier $F[f]$: el valor de la funcional $F[f]$ en todo punto φ del espacio S es igual al valor de la funcional f en el punto $F[\varphi] \in S$.

Como ejemplo, hallemos la transformación de Fourier de la unidad, considerándola como una función generalizada. Evidentemente, $1 \in S'$. Tenemos

$$(\hat{1}, \varphi) = (1, \hat{\varphi}) = \int_{-\infty}^{+\infty} dy \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{-ixy}\, dx =$$

$$= \sqrt{2\pi} \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} \varphi(x) e^{iy(t-x)}\, dx \Big|_{t=0} =$$

$$= \sqrt{2\pi}\, \varphi(t)\Big|_{t=0} = \sqrt{2\pi}\, (\delta, \varphi)$$

(aquí hemos aprovechado el lema 1 del p. 56.5). De este modo, $1 = \sqrt{2\pi}\delta$.

Observemos que la transformación de Fourier $F[\varphi]$ de la función $\varphi \in D$, no pertenece, en el caso general, al espacio D, por cuanto $F[\varphi]$ no es siempre una función finita. Por eso la fórmula (59.32) tiene sentido no para todas las $f \in D'$. Precisamente debido a esta circunstancia, al considerar la transformación de Fourier de las funciones generalizadas, tuvimos que hacer más estrecha la clase de funciones generalizadas introducidas anteriormente, limitándonos sólo a las funciones generalizadas de crecimiento lento.

La transformación de Fourier $F[f]$ de una función generalizada f se designará por el símbolo $\hat{f}$ o el símbolo

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ixy}\, dx.$$

De este modo, la igualdad

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ixy}\, dx = F[f] \tag{59.33}$$

es, cuando f es una función generalizada, la definición del símbolo que figura en el primer miembro de esta igualdad.

Al definir la transformación de Fourier para todas las funciones generalizadas de S', hemos definido también, en particular, la transformación de Fourier para las funciones ordinarias f que satisfacen la condición (59.25), es decir, para las funciones de una clase mucho más amplia en comparación con lo hecho anteriormente (véase el p. 56.5 y 58.7*). Esto constituye una circunstancia más que justifica la conveniencia de introducción del concepto de funciones generalizadas.

Mostremos que la transformación de Fourier de las funciones generalizadas posee toda una serie de propiedades análogas a las de la transformación de Fourier clásica, es decir, de la transformación de Fourier de las funciones absolutamente integrables.

Lema 7. *La transformación de Fourier $F[f]$ de una función generalizada $f \in S'$ es también una función generalizada de la clase S', es decir, $F[f]$ es una funcional continua y lineal sobre el espacio S.*

DEMOSTRACIÓN. Comprobemos la linealidad de la transformación de Fourier, es decir, probemos que, cualquiera que sea la función generalizada $f \in S'$, para cualesquiera funciones $\varphi \in S$, $\psi \in S$ y cualesquiera números λ y μ se verifica la igualdad

$$(F[f], \lambda\varphi + \mu\psi) = \lambda(F[f], \varphi) + \mu(F[f], \psi).$$

En efecto,

$$(F[f], \lambda\varphi + \mu\psi) = (f, F[\lambda\varphi + \mu\psi]) =$$
$$= (f, \lambda F[\varphi] + \mu F[\psi]) = \lambda(f, F[\varphi]) + \mu(f, F[\psi]) =$$
$$= \lambda(F[f], \varphi) + \mu(F[f], \psi).$$

Comprobemos la continuidad de la transformación de Fourier. Sea $f \in S'$, $\varphi \in S$, $\varphi_n \in S$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} \varphi_n = \varphi$, y por lo tanto (véase el teorema 1 del p. 59.6),

$$\lim_{n \to \infty} F[\varphi_n] = F[\varphi].$$

Entonces, por ser la funcional f continua en S, obtenemos

$$\lim_{n \to \infty} (F[f], \varphi_n) = \lim_{n \to \infty} (f, F[\varphi_n]) = (f, F[\varphi]) = (F[f], \varphi).$$

Hemos probado, pues que si $f \in S'$, entonces también $F[f] \in S'$. □

Del modo natural se define también la transformación de Fourier inversa $F^{-1}[f]$ del elemento $f \in S'$, como una funcional del espacio S', definida por la fórmula

$$(F^{-1}[f], \varphi) = (f, F^{-1}[\varphi]), \quad \varphi \in S.$$

Si f es una función absolutamente integrable, dicha igualdad se verifica para esta función en el sentido corriente. Esto se comprueba igual que en el caso de la fórmula (59.31). Por definición, se presupone también que (compárese con (59.33))

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{ixy}\, dx = F^{-1}[f]. \tag{59.34}$$

Lo mismo que en el caso de una transformación de Fourier directa F, se muestra que si $f \in S'$, entonces también $F^{-1}[f] \in S'$.

Teorema 2. *La transformación de Fourier F y la transformación de Fourier inversa F^{-1} aplican lineal, biunívoca y continuamente el espacio S' sobre sí mismo; además, para cualquier elemento $f \in S'$ se verifican las igualdades*

$$F^{-1}[F[f]] = F[F^{-1}[f]] = f. \tag{59.35}$$

DEMOSTRACIÓN. Demostremos primero las fórmulas (59.35). Para cualquier elemento $\varphi \in S$ se tiene

$$(F^{-1}[F[f]], \varphi) = (F[f], F^{-1}[\varphi]) = (f, F[F^{-1}[\varphi]]) = (f, \varphi).$$

Por analogía,

$$(F[F^{-1}[f]], \varphi) = (F^{-1}[f], F[\varphi]) = (f, F^{-1}[F[\varphi]]) = (f, \varphi).$$

Mostremos ahora que la transformación de Fourier F aplica el espacio S' sobre todo el espacio S': $F(S') = S'$. Sea $g \in S'$, entonces si $f = F^{-1}[g]$, se tiene $F[f] = F[F^{-1}[g]] = g$, es decir, realizándose la transformación de Fourier F, en cualquier elemento de S' se aplica cierto elemento de S'.

Mostremos que F es biunívoca. Si $f_1 \in S'$, $f_2 \in S'$ y $F[f_1] = F[f_2]$, entonces también $F^{-1}[F[f_1]] = F^{-1}[F[f_2]]$, de donde, en virtud de (59.35), se tiene $f_1 = f_2$.

Señalemos que la aplicación F es lineal, es decir, que para cualesquiera funciones generalizadas $f \in S'$, $g \in S'$ y cualesquiera números λ y μ se verifica la igualdad

$$F[\lambda f + \mu g] = \lambda F[f] + \mu F[g].$$

Con el fin de convencernos de la validez de esta igualdad, comprobémosla para cualquier elemento $\varphi \in S$, siempre que sea éste fijo.

$$(F[\lambda f + \mu g], \varphi) = (\lambda f + \mu g, F[\varphi]) = \lambda(f, F[\varphi]) + \mu(g, F[\varphi]) =$$

$$= \lambda(F[f], \varphi) + \mu(F[g], \varphi) = (\lambda F[f] + \mu F[g], \varphi).$$

Por fin, demostremos que F es una aplicación continua. En efecto, sea $f \in S'$, $f_n \in S'$, $n = 1, 2, \ldots$, $\lim f_n = f$, y, por consiguiente, para cualquier $\varphi \in S$ se verifica la igualdad $\lim_{n \to \infty} (f_n, \varphi) = (f, \varphi)$. En tal caso

$$\lim_{n \to \infty} (F[f_n], \varphi) = \lim_{n \to \infty} (f_n F[\varphi]) = (f, F[\varphi]) = (F[f], \varphi).$$

Análogamente se demuestra que también F^{-1} aplica continua y biunívocamente S' sobre S'. □

Ejemplos. Hallemos $F[\delta] = \hat{\delta}$. Tenemos

$$(\hat{\delta}, \varphi) = (\delta, \hat{\varphi}) = \hat{\varphi}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{-ixy}\, dx \Big|_{y=0} =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x)\, dx = \left(\frac{1}{\sqrt{2\pi}}, \varphi\right), \quad \varphi \in S,$$

por lo cual $F[\delta] = \frac{1}{\sqrt{2\pi}}$, y por lo tanto, $F^{-1}[1] = \sqrt{2\pi}\delta$ (observemos que la transformación de Fourier inversa clásica $F^{-1}[1]$, al igual que la transformación directa $F[1]$, no existen). Con ayuda de las integrales (59.33) y (59.34) estas fórmulas

pueden escribirse en la forma

$$\int_{-\infty}^{+\infty} \delta(x)e^{-ixy}\,dx = 1, \quad \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ixy}\,dx = \delta(y).$$

De modo igual se halla también la transformación de Fourier inversa de la función delta:

$$F^{-1}[\delta] = \frac{1}{\sqrt{2\pi}} = F[\delta],$$

de donde

$$F[1] = F^{-1}[1] = \sqrt{2\pi}\,\delta.$$

Al hacer uso de la notación basada sobre las igualdades (59.33) y (59.34), estas fórmulas pueden escribirse en la forma

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ixy}\,dx = \delta(y), \quad \int_{-\infty}^{+\infty} \delta(x)e^{ixy}\,dx = 1.$$

Calculemos ahora la transformación de Fourier de la derivada de una función generalizada y la derivada de la transformación de Fourier. Previamente hemos de introducir el concepto de producto de una función generalizada $f \in S'$ por una función habitual infinitamente derivable $\psi(x)$ que posee la propiedad de que para cualquiera de sus derivadas $\psi^{(n)}(x)$ existen unas constantes $\beta_n > 0$ y $\alpha_n > 0$, $n = 0, 1, \ldots$, tales que para todos los x se verifica la desigualdad

$$|\psi^{(n)}(x)| \leqslant \beta_n(1 + |x|)^{\alpha_n}, \quad n = 0, 1, 2, \ldots .^{*)} \tag{59.36}$$

Señalemos que todos los polinomios satisfacen esta condición.

Si la función ψ es del tipo (59.36) y $\varphi \in S$, entonces $\psi\varphi \in S$. Si la función f es localmente sumable y satisface la condición (59.25), mientras que la función ψ satisface la condición (59.36), entonces ψf satisface también la condición (59.25) y

$$(f, \psi\varphi) = \int_{-\infty}^{+\infty} f(x)\psi(x)\varphi(x)\,dx = (\psi f, \varphi).$$

Supongamos que ψ satisface la condición (59.36) y $f \in S'$. Definamos ahora la funcional sobre S, igual al producto ψf, mediante la fórmula

$$(\psi f, \varphi) = (f, \psi\varphi), \quad \varphi \in S.$$

Es fácil comprobar que $\psi f \in S'$ **), es decir, ψf es una funcional lineal continua de-

*) En virtud de esta condición (para $n = 0$), $\psi(x)$ puede considerarse como una función generalizada del espacio S' (véase (59.25)).

**) Las complicaciones que aparecen al determinar el producto de las funciones generalizadas se deben a que el producto de las funcionales lineales en el sentido habitual como un producto de funciones (es decir, como producto de los valores de factores en todo punto) no es una funcional lineal.

finida sobre el espacio S.

Ejercicio 21. Supongamos que la función $\psi = \psi(x)$ satisface la condición (59.36) y la función generalizada $f \in S'$. Demuéstrese que $\psi f \in S$.

Demostraremos en conclusión las fórmulas

$$F[f^{(n)}] = (ix)^n F[f], \tag{59.37}$$

$$i^n F^{(n)}[f] = F[x^n f], \quad f \in S. \tag{59.38}$$

Tenemos (véase el p. 56.8):

$$(F[f^{(n)}], \varphi) = (f^{(n)}, F[\varphi]) = (-1)^n (f, F^{(n)}[\varphi]) =$$

$$= (-1)^n \left(f, \frac{1}{i^n} F[x^n \varphi]\right) = i^n (F[f], x^n \varphi) = ((ix)^n F[f], \varphi), \quad \varphi \in S.$$

La fórmula (59.37) está demostrada.

Demostremos (59.38) (véase el p. 56.10):

$$(F^{(n)}[f], \varphi) = (-1)^n (F[f], \varphi^{(n)}) = (-1)^n (f, F[\varphi^{(n)}]) =$$

$$= (-1)^n (f, (ix)^n F[\varphi]) = \frac{1}{i^n} (x^n f, F[\varphi]) = \left(\frac{1}{i^n} F[x^n f], \varphi\right). \square$$

Ejemplo. Hallemos la transformación de Fourier de la función $f(x) = x$:

$$F[x] = F[x \cdot 1] = iF'[1] = i\sqrt{2\pi}\delta'$$

Ejercicio 22. Hállese la transformación de Fourier de un polinomio.

Al calcular una transformación de Fourier de las funciones generalizadas resulta a veces conveniente elegir una sucesión de funciones ordinarias que en el espacio S' tienden a una función (generalizada) dada, hallar la transformación de Fourier de los términos de dicha sucesión y después calcular la transformación de Fourier buscada de la función dada pasando al límite y aprovechando la continuidad de la transformación de Fourier. Así por ejemplo, para calcular la transformación de Fourier $F[\theta]$ de la función de Heaviside $\theta(x)$, hallemos primero la transformación de Fourier de la función $\theta(x)e^{-tx}$ $(t > 0)$.

$$F[\theta(x)e^{-tx}] = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-x(t+iy)}\, dx = -\frac{e^{-x(t+iy)}}{\sqrt{2\pi}(t+iy)} \Bigg|_0^{+\infty} =$$

$$= \frac{1}{\sqrt{2\pi}(t+iy)} = -\frac{i}{\sqrt{2\pi}(y-it)}. \tag{59.39}$$

Probemos ahora que en S'

$$\lim_{t \to +0} \theta(x)e^{-tx} = \theta(x) \tag{59.40}$$

En efecto, para toda función $\varphi \in S$ y cualquier número A se tiene

$$|(\theta(x), \varphi(x)) - (\theta(x)e^{-tx}, \varphi(x))| = \left| \int_0^{+\infty} (1 - e^{-tx})\varphi(x)\, dx \right| \leqslant$$

$$\leqslant \left| \int_0^A (1 - e^{-tx})\varphi(x)\, dx \right| + \left| \int_A^{+\infty} (1 - e^{-tx})\varphi(x)\, dx \right|. \tag{59.14}$$

Fijemos la función $\varphi \in S$ y un número cualquiera $\varepsilon > 0$. Por ser la función φ absolutamente integrable, existe un número $A > 0$ tal que

$$\int_A^{+\infty} |\varphi(x)|\, dx < \frac{\varepsilon}{2};$$

entonces

$$\left| \int_A^{+\infty} (1 - e^{-tx})\varphi(x)\, dx \right| \leqslant \int_A^{+\infty} |\varphi(x)|\, dx < \frac{\varepsilon}{2}. \tag{59.42}$$

Elijamos ahora $t_0 > 0$ de modo tal que para $0 < t < t_0$ sea válida la desigualdad

$$(1 - e^{-tA}) \int_0^A |\varphi(x)|\, dx < \frac{\varepsilon}{2}$$

y, por consiguiente,

$$\left| \int_0^A (1 - e^{-tx})\varphi(x)\, dx \right| < (1 - e^{-tA}) \int_0^A |\varphi(x)|\, dx < \frac{\varepsilon}{2}. \tag{59.43}$$

Entonces, para $0 < t < t_0$, de (59.41), (59.42) y (59.43) obtenemos

$$|(\theta(x), \varphi(x)) - (\theta(x)e^{-tx}, \varphi(x))| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

La fórmula (59.40) queda demostrada.

Por ser la transformación de Fourier continua, tenemos

$$\lim_{t \to +0} F[\theta(x)e^{-tx}] = F[\theta(x)]; \tag{59.44}$$

de aquí y de (59.39) tenemos

$$F[\theta(x)] = -\frac{1}{\sqrt{2\pi}} \lim_{t \to +0} \frac{i}{y - it},$$

con la particularidad de que de (59.44) proviene que el límite en el segundo miembro existe (en el espacio S'), este límite se designa, corrientemente, por $\frac{i}{y - i0}$ (véase el ejercicio 8).

De este modo,

$$F[\theta(x)] = -\frac{1}{\sqrt{2\pi}} \frac{i}{y - i0}.$$

Ejercicio 23. Hállese la transformación de Fourier de las funciones $x^k\theta(x)$, $k = 1, 2, \ldots$.

COMPLEMENTO

§ 60. ALGUNOS PROBLEMAS DE LOS CÁLCULOS APROXIMADOS

60.1. APLICACIÓN DE LA FÓRMULA DE TAYLOR PARA EL CÁLCULO APROXIMADO DE LOS VALORES DE FUNCIONES E INTEGRALES

Para calcular los valores de las funciones resulta cómodo servirse de la fórmula de Taylor o serie de Taylor. Aclaremos esto con los ejemplos.

1. Cálculo del valor del seno.

La fórmula de Taylor para la función sen x tiene por expresión

$$\operatorname{sen} x = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} + r_n(x),$$

donde

$$r_n(x) = (-1)^n \frac{x^{2n+1}}{(2n+1)!} \operatorname{sen}^{(2n+1)} \theta x, \quad 0 < \theta < 1$$

(el término residual se ha tomado en la forma de Lagrange). Por esto

$$|r_n(x)| < \frac{|x|^{2n+1}}{(2n+1)!}. \tag{60.1}$$

Supongamos que se pide hallar sen 20° con un error inferior a 10^{-3}. A 20° medidos en radianes corresponde la magnitud $\frac{\pi}{9}$, por lo cual elijamos el número n de modo tal que sea

$$\left| r_n\left(\frac{\pi}{9}\right) \right| < \frac{1}{10^3}; \tag{60.2}$$

entonces el valor del polinomio de Taylor de orden n en el punto $x = \frac{\pi}{9}$ nos dará la aproximación buscada de sen 20°. En virtud de la desigualdad (60.1), para que se cumpla la condición (60.2), es suficiente que se verifique la desigualdad

$$\frac{1}{(2n+1)!} \left(\frac{\pi}{9}\right)^{2n+1} < \frac{1}{10^3}. \tag{60.3}$$

Cuando $n = 1$, esta desigualdad no se verifica:

$$\frac{1}{3!}\left(\frac{\pi}{9}\right)^3 > \frac{1}{6}\,\frac{1}{3^3} = \frac{1}{162} > \frac{1}{10^3},$$

pero, cuando $n = 2$, sí se cumple:

$$\frac{1}{5!}\left(\frac{\pi}{9}\right)^5 < \frac{1}{120}\cdot\frac{1}{2^5} = \frac{1}{3840} < \frac{1}{10^3}.$$

Por eso, sen 20° con un error inferior a 10^{-3} se halla según la fórmula

$$\operatorname{sen} 20^\circ \approx \frac{\pi}{9} - \frac{1}{6}\left(\frac{\pi}{9}\right)^3 \text{*)}. \tag{60.4}$$

Si tomamos π de las tablas con un error inferior a 10^{-4}, sustituimos este valor en la fórmula (60.4), realizamos las operaciones dictadas por la fórmula y redondeamos el resultado con exactitud hasta de 10^{-3}, obtendremos la aproximación buscada de sen 20°:

$$\operatorname{sen} 20^\circ \approx 0{,}343 \text{**)}.$$

Calculando los valores de un seno se puede emplear, en lugar de la fórmula de Taylor, una serie de Taylor la que es alternada para el argumento real y admite, por esta razón, una estimación sencilla del resto: no es superior, en valor absoluto, al valor absoluto del primer término del resto (véase el p. 35.9). Esto nos da, naturalmente, el mismo resultado que antes, puesto que conduce a la estimación (60.3) la cual se ha obtenido partiendo de otros razonamientos.

2. Cálculo de los valores de logaritmos naturales.

La serie de Taylor para un logaritmo

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{x^n}{n}, \quad -1 < x \leqslant 1, \tag{60.5}$$

puede ser utilizado de inmediato sólo para calcular logaritmos de los números que no sobrepasan dos. Sin embargo, de la serie (60.5) pueden obtenerse otros desarrollos que permiten calcular los logaritmos de cualesquiera números. Al sustituir en (60.5) x por $-x$ y sustraer la serie obtenida de (60.5), obtendremos

---

*) Mediante el signo $\approx$ se denota la igualdad aproximada con el grado prefijado de exactitud.

**) Señalemos que en nuestro caso se establece con facilidad una desigualdad más fuerte aún: $r_2\left(\frac{\pi}{9}\right) < \frac{1}{3}\,10^{-3}$, mientras que siendo prefijado el número de signos de π, el error, que se obtiene al calcular el segundo miembro de la fórmula (60.4), en todo caso no sobrepasa $\frac{2}{3}\cdot 10^{-3}$, por lo cual el error total será no superior a 10^{-3}.

$$\ln\frac{1+x}{1-x} = 2x\sum_{n=0}^{\infty}\frac{x^{2n}}{2n+1}, \quad |x| < 1. \tag{60.6}$$

Cuando x varía desde -1 hasta 1, la expresión $\dfrac{1+x}{1-x}$ toma todos los valores positivos. Por ello, la fórmula (60.6) puede utilizarse para calcular logaritmos de cualesquiera números positivos. Surge naturalmente una pregunta ¿cuántos términos se deben tomar en la serie (60.6), para que se obtenga el logaritmo de un número con la exactitud dada? Con este fin es preciso estimar el resto de la serie (60.6). Tenemos

$$|r_n(x)| = 2|x|\sum_{k=n}^{\infty}\frac{x^{2k}}{2k+1} < \frac{2|x|^{2n+1}}{2n+1}\sum_{k=0}^{\infty}x^{2k} = \frac{2|x|^{2n+1}}{(2n+1)(1-x^2)}, \quad |x| < 1. \tag{60.7}$$

Hagamos uso de esta estimación para calcular ln 2 con un error inferior a 10^{-3}. Resolviendo la ecuación

$$\frac{1+x}{1-x} = 2,$$

encontramos $x = \dfrac{1}{3}$. Haciendo en (60.6) $x = \dfrac{1}{3}$, encontramos

$$\ln 2 = \frac{2}{3}\sum_{n=0}^{\infty}\frac{1}{(2n+1)3^{2n}}. \tag{60.8}$$

La estimación (60.7) nos da en este caso

$$\left|r_n\left(\frac{1}{3}\right)\right| < \frac{2}{(2n+1)3^{2n+1}}\cdot\frac{1}{1-\frac{1}{9}} = \frac{1}{4(2n+1)3^{2n-1}}.$$

De aquí, para $n = 3$, tenemos

$$r_3\left(\frac{1}{3}\right) < \frac{1}{4\cdot 7\cdot 3^5} = \frac{1}{28\cdot 243} < \frac{1}{10^3}.$$

Por ello, para calcular ln 2 con un error inferior a 10^{-3}, es suficiente tomar los primeros tres términos de la serie (60.8):

$$\ln 2 \approx \frac{2}{3}\left(1 + \frac{1}{3^3} + \frac{1}{5\cdot 3^4}\right) \approx 0{,}693.$$

Al calcular los valores de una función con menor exactitud, sirviéndose de la fórmula de Taylor

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x),$$

es a menudo suficiente limitarse sólo a la parte lineal, es decir, a los primeros dos términos

$$f(x) = f(x_0) + f'(x_0)(x - x_0),$$

en otras palabras, sustituir el incremento de la función por la diferencial de ésta

$$\Delta y = f(x) - f(x_0) \approx f'(x_0)(x - x_0) = f'(x_0)\Delta x,$$

donde $\Delta x = x - x_0$.

La fórmula de Taylor permite también calcular aproximadamente los valores de ciertas integrales. Examinemos un ejemplo de esta índole.

3. Cálculo de la integral

$$\int_0^1 \frac{\operatorname{sen} x}{x}\,dx.$$

con un error inferior a 0,0001.

Escribamos la fórmula de Taylor para el integrando. Hagamos uso con este fin de la fórmula conocida de Taylor para la función sen x (véase (60.1)), en este caso obtendremos

$$\frac{\operatorname{sen} x}{x} = \sum_{k=1}^{n} (-1)^{k-1}\frac{x^{2k-2}}{(2k-1)!} + \frac{r_n(x)}{x};$$

por eso

$$\int_0^1 \frac{\operatorname{sen} x}{x}\,dx = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{(2k-1)!}\int_0^1 x^{2k-2}dx + \int_0^1 \frac{r_n(x)}{x}\,dx.$$

Teniendo presente la estimación (60.1), resulta

$$\left|\int_0^1 \frac{r_n(x)}{x}\,dx\right| \leqslant \int_0^1 \frac{|r_n(x)|}{x}\,dx \leqslant \frac{1}{(2n+1)!}\int_0^1 x^{2n}dx = \frac{1}{(2n+1)!(2n+1)}.$$

Por cuanto, para $n = 3$,

$$\frac{1}{(2n+1)!(2n+1)} = \frac{1}{7!7} = \frac{1}{35280} < \frac{1}{3}10^{-4},$$

entonces, con un error inferior a 0,0001 tenemos

$$\int_0^1 \frac{\operatorname{sen} x}{x}\,dx \approx \int_0^1 dx - \frac{1}{6}\int_0^1 x^2dx + \frac{1}{120}\int_0^1 x^4dx = 1 - \frac{1}{18} + \frac{1}{600} \approx 0{,}9961^{*}).$$

*) Al convertir las fracciones simples en las decimales se ha cometido un error no superior a $\frac{1}{2}\cdot 10^{-4}$, por lo cual el error total, realizado el cálculo aproximado de la integral en consideración, realmente no sobrepasa la magnitud de 10^{-4}.

Observemos que en los cálculos aproximados prácticos de las integrales la fórmula de Taylor resulta ser, por regla general, inconveniente para emplearla, por cuanto en dicha fórmula figuran las derivadas de la función dada y el cálculo de éstas conduce a la acumulación complementaria de errorres. Resulta más oportuno emplear las fórmulas aproximadas de integración en las que figuran sólo los valores de la propia función. Los métodos semejantes de la integración aproximada se considerarán en el p. 60.4.

OBSERVACIÓN. Cabe advertir que para cálculos reales de los valores de funciones o de integrales de dichas funciones, desarrollando las funciones en series, no todo desarrollo de las funciones consideradas en series será aplicable. Puede suceder que la serie obtenida convergerá de modo tan "lento" que prácticamente o bien no será útil para los cálculos o bien estos últimos serán injustificadamente engorrosos (hablando metafóricamente, en este caso la serie es "prácticamente divergente", aunque "converge teóricamente"). En tal situación se debe tratar de obtener alguna otra serie que converja lo suficientemente rápido ("mejorar la convergencia", como suele decirse en este caso) y cuya suma permitirá hallar los valores de la función en consideración. Precisamente así hemos procedido al considerar el método para calcular logaritmos. Sería, por ejemplo, inconveniente calcular incluso el valor de $\ln \frac{3}{2}$ con ayuda de la serie (60.5), aunque la serie converge cuando $x = \frac{1}{2}$; en lugar de esto se debe utilizar de la serie (60.6) para $x = \frac{1}{5}$, puesto que dicha serie converge más rápidamente.

60.2. RESOLUCIÓN DE LAS ECUACIONES

Examinemos un problema en el que se resuelve la ecuación

$$f(x) = 0. \tag{60.9}$$

Si la función f es continua en el segmento $[a, b]$ y toma en los extremos del segmento los valores de signo opuesto, entonces el método, por cuyo intermedio se ha demostrado en el p. 6.2 el teorema sobre la existencia en este caso de un punto x_0, donde la función se anula, nos proporciona también el procedimiento para calcular dicho valor x_0, es decir, la raíz de la ecuación (60.9). Con este fin es suficiente dividir sucesivamente el segmento $[a, b]$ en dos partes iguales, eligiendo cada vez aquel segmento en cuyos extremos la función f adquiere valores de signo contrario (siempre que, por supuesto, no ocurre que en uno de los extremos obtenidos la función f se anule, caso en el cual la raíz buscada ya se habrá encontrado). Si se pide hallar la raíz de la ecuación (60.9) con un error inferior a $\varepsilon > 0$ dado, entonces, realizados n pasos tales que

$$\frac{b-a}{2^n} < \varepsilon,$$

los extremos del segmento obtenido nos darán precisamente la aproximación buscada de cierta raíz de la ecuación (60.9) (el extremo izquierdo por defecto y el derecho

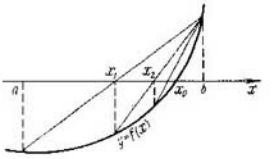


Fig. 246

por exceso). Este procedimiento de resolución aproximada de la ecuación (60.9), llamado "método de horquilla", es, en principio, muy simple, aunque bastante voluminoso. Se emplea principalmente para el "tanteo aproximado" del resultado, o sea, para la determinación "aproximada" del intervalo, en el que se halla la raíz buscada de la ecuación que se resuelve, y a continuación, con el objeto de encontrar en dicho intervalo un valor "más exacto" de la raíz se utilizan otros métodos que convergen con mayor rapidez: se aplica, corrientemente, el método de tangentes ("método de Newton") que se describe más abajo. Por regla general, tal esquema se aplica al realizar los cálculos en las máquinas calculadoras de accionamiento rápido. Naturalmente, este mismo procedimiento resulta provechoso también al calcular "a mano", en particular, con ayuda de una regla de cálculo o minicomputadores.

Analizaremos aquí los métodos de resolución de las ecuaciones que se denominan *método de las cuerdas y método de las tangentes*. El último método se generaliza con éxito al caso de los sistemas de ecuaciones.

En lo que sigue supondremos siempre que la función f es continua en el segmento $[a, b]$ y tiene en dicho segmento las derivadas primera y segunda*), con la particularidad de que ambas son de signo constante (en particular, distintas de cero).

Supondremos, además, que la función f toma en los extremos del segmento valores de signo opuesto. Ya que la primera derivada es de signo constante, la función f es estrictamente monótona, razón por la cual bajo las suposiciones asumidas la ecuación (60.9) tiene en el intervalo (a, b) exactamente una raíz.

MÉTODO DE LAS CUERDAS

Este método consiste en lo siguiente. La gráfica de la función f se sustituye por su cuerda, es decir, por un segmento que une los puntos extremos de la gráfica de la función f: los puntos $(a, f(a))$ y $(b, f(b))$. La abscisa x_1 del punto de intersección de esta cuerda con el eje Ox se considera como la primera aproximación de la raíz buscada (fig. 246). Luego se toma aquel de los segmentos $[a, x_1]$ y $[x_1, b]$, en cuyos extremos la función f adquiere valores de signo opuesto (en adelante se mostrará que bajo las suposiciones asumidas $f(x_1) \neq 0$, y, por consiguiente, tal segmento

*) Para el método de cuerdas basta exigir que las derivadas primera y segunda existan sólo en el intervalo (a, b). La existencia de la derivada en los extremos del segmento $[a, b]$ se utilizará sólo en el método de las tangentes.

siempre existe), y a dicho segmento se aplica el mismo procedimiento; se obtiene la segunda aproximación de la raíz x_2, etc. De resultas, se forma una sucesión x_n, $n = 1, 2, \ldots$, la cual (como se probará más abajo) converge, con las restricciones impuestas sobre la función f, hacia la raíz de la ecuación (60.9).

No es difícil obtener las fórmulas recurrentes para los números citados x_n, $n = 1, 2, \ldots$. La ecuación de una recta que pasa por los puntos extremos de la gráfica de la función f tiene por expresión

$$y = \frac{f(b) - f(a)}{b - a}(x - a) + f(a). \tag{60.10}$$

Designemos su miembro derecho mediante $l(x)$, es decir, escribamos la ecuación (60.10) en la forma

$$y = l(x).$$

Hallemos la abscisa x_1 del punto de intersección de la recta (60.10) con el eje Ox, es decir, resolvamos la ecuación $l(x) = 0$; obtendremos

$$x_1 = a - \frac{(b - a)f(a)}{f(b) - f(a)}. \tag{60.11}$$

Es fácil convencerse de que

$$a < x_1 < b \tag{60.12}$$

(esto se deduce, por ejemplo, del hecho de que la función $l(x)$ es estrictamente monótona y continua y de que en los extremos del segmento $[a, b]$ la función toma valores de signo opuesto: $l(a) = f(a)$ y $l(b) = f(b)$).

Análogamente encontramos

$$x_{n+1} = x_n - \frac{(b - x_n)f(x_n)}{f(b) - f(x_n)}, n = 1, 2, \ldots. \tag{60.13}$$

Mostremos que la sucesión $\{x_n\}$ tiende a la raíz de la ecuación (60.9) de manera monótona. Supongamos, para concretar, que $f'(x) > 0, f''(x) > 0, a < x < b$ (véase la fig. 246). En este caso la función f es estrictamente creciente y gira su convexidad hacia las y negativas. Por consiguiente cualquier punto interior de la cuerda, que une los puntos extremos de la gráfica de la función f, se sitúa por encima del punto correspondiente de la gráfica de la función f, es decir,

$$l(x) > f(x), a < x < b.$$

En particular, si x_0 es la raíz de la ecuación (60.9): $f(x_0) = 0$, entonces de aquí proviene que

$$l(x_0) > 0.$$

Tenemos (véanse (60.11) y (60.12)):

$$l(x_1) = 0, a < x_1 < b.$$

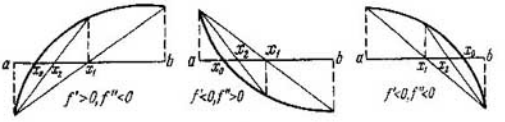


Fig. 247

De este modo,

$$l(x_1) < l(x_0), \tag{60.14}$$

pero la función lineal $l(x)$ crece de modo estrictamente monótono, pues

$$l(b) = f(b) > f(a) = l(a),$$

por lo cual de (60.14) se deduce que

$$x_1 < x_0.$$

Ahora, al sustituir el segmento $[a, b]$ por el segmento $[x_1, b]$ y observar que $f(x_1) < 0$, demostremos análogamente que

$$x_1 < x_2 < x_0.$$

Luego, por inducción, obtendremos

$$x_1 < x_2 < \ldots < x_n \ldots < x_0.$$

Así pues, la sucesión $\{x_n\}$ converge, puesto que es monótona y acotada. Sea $\lim_{n \to \infty} x_n = c$. Pasando en la igualdad (60.13) al límite para $n \to \infty$, obtendremos $f(c) = 0$, es decir, la sucesión $\{x_n\}$ converge hacia la raíz de la ecuación (60.9).

Si $|f'(x)| \geqslant m > 0$, $a < x < b$, entonces no es difícil de obtener la estimación de la velocidad con la cual converge la sucesión $\{x_n\}$ en términos de la propia función f en los puntos x_n. En efecto,

$$f(x_n) = f(x_n) - f(x_0) = f'(\xi_n)(x_n - x_0),$$
$$x_n < \xi_n < x_0, \; n = 1, 2, \ldots;$$

de aquí

$$|x_n - x_0| \leqslant \frac{|f(x_n)|}{m}, \; n = 1, 2, \ldots.$$

Los casos restantes, a saber,

$$f'(x) > 0, f''(x) < 0,$$
$$f'(x) < 0, f''(x) > 0,$$
$$f'(x) < 0, f''(x) < 0,$$

se consideran por analogía con el caso que acabamos de examinar (fig. 247).

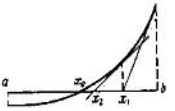


Fig. 248

MÉTODO DE LAS TANGENTES (MÉTODO DE NEWTON)

Supondremos que la función f satisface las mismas condiciones que eran vigentes al considerar el método de las cuerdas. Tracemos una tangente a la gráfica de la función f en uno de sus puntos extremos, por ejemplo, en el punto $(b, f(b))$. La abscisa x_1 del punto de intersección de la tangente con el eje Ox se considera precisamente la primera aproximación de la raíz de la ecuación (60.9). Luego, si $x_1 \in (a, b)$ (según lo expuesto más abajo, esto siempre tiene lugar para una de las tangentes en los puntos extremos de la gráfica), entonces de los dos segmentos $[a, x_1]$ y $[x_1, b]$ se elige aquel, en cuyos extremos la función f toma los valores de signo opuesto (en adelante se mostrará que $f(x_1) \neq 0$). A continuación se traza la tangente a la gráfica de la función f en el punto $(x_1, f(x_1))$; el punto de su intersección con el eje Ox se denota con x_2, etc. (fig. 248).

Se obtienen con facilidad fórmulas recurrentes para los números citados x_n, $n = 1, 2, \ldots$. La ecuación de la tangente que pasa por el punto $(b, f(b))$ tiene por expresión

$$y = f'(b)(x - b) + f(b).$$

Designemos su miembro derecho mediante $L(x)$, es decir, escribamos esta ecuación en la forma

$$y = L(x).$$

Hallemos la abscisa x_1 del punto de intersección de esta tangente con el eje Ox, es decir, resolvamos la ecuación $L(x) = 0$; se obtendrá:

$$x_1 = b - \frac{f(b)}{f'(b)}.$$

El punto x_1 puede disponerse, en el caso general, fuera del segmento $[a, b]$, es decir, fuera del dominio de definición de la función f. Sin embargo, si $f(b)$ y f'' son de un mismo signo, entonces $x_1 \in (a, b)$. Igual que al describir el método de las cuerdas, examinaremos más detalladamente el caso en que $f' > 0, f'' > 0$ en $[a, b]$. En este caso la función f crece de manera estrictamente monótona, por consiguiente, $f(b) > 0$; además, la función f es convexa hacia las y negativas en (a, b), por consiguiente,

$$L(x) < f(x)$$

(véase el p. 14.3).

Si $f(x_0) = 0$, $a < x_0 < b$, se tiene

$$L(x_0) < 0,$$

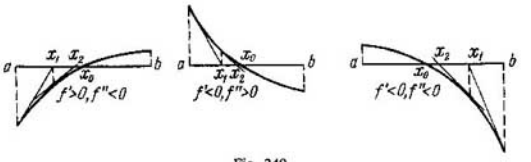


Fig. 249

pero $L(b) = f(b) > 0$, por consiguiente,

$$x_0 < x_1 < b.$$

Con ello, $f(x_1) > L(x_1) = 0$.

Razonando análogamente respecto del segmento $[a, x_1]$, obtendremos un punto x_2 tal que

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}, \quad x_0 < x_2 < x_1,$$

y, luego,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 < x_{n+1} < x_n. \tag{60.15}$$

Por consiguiente, la sucesión $\{x_n\}$ es monótona y acotada y por esta razón converge. Sea $\lim_{n \to \infty} x_n = c$. Pasando al límite en (60.15), obtendremos $f(c) = 0$, es decir, la sucesión (60.15) converge hacia la raíz $x_0 = c$ de la ecuación (60.9).

Cuando $|f'(x)| \geqslant m \geqslant 0$, $a < x < b$, por el procedimiento utilizado en el método de las cuerdas, obtendremos la estimación

$$|x_n - c| \leqslant \frac{|f(x_n)|}{m}, \quad n = 1, 2, \ldots .$$

De un modo semejante se examinan también los demás casos de combinaciones diferentes de los signos que tienen las derivadas primera y segunda (fig. 249).

He aquí una estimación más de la velocidad de convergencia del método de las tangentes, de la que se ve claramente la ventaja del método mencionado. Supongamos que para la función f en el intervalo que se considera se cumplen las desigualdades

$$|f'(x)| \geqslant m > 0, \quad |f''(x)| \leqslant M, \quad a < x < b.$$

Desarrollemos la función f en el entorno del punto x_n según la fórmula de Taylor, por ejemplo, con el término residual en la forma de Lagrange

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{1}{2} f''(\xi)(x - x_n)^2,$$

donde $\xi = x_n + \theta(x - x_n)$, $0 < \theta < 1$. Si $f(c) = 0$, entonces, al sustituir $x = c$ en la fórmula escrita, obtendremos

$$f(x_n) + f'(X_n)(c - x_n) + \frac{1}{2} f''(\xi)(c - x_n)^2 = 0.$$

De aquí

$$x_n - c = \frac{f(x_n)}{f'(x_n)} + \frac{f''(\xi)}{2f'(x_n)} (c - x_n)^2,$$

o bien, en virtud de la fórmula (60.15),

$$|x_{n+1} - c| = \frac{f''(\xi)}{2f'(x_n)} (c - x_n)^2.$$

Por consiguiente,

$$|x_{n+1} - c| \leqslant \frac{M}{2m} |x_n - c|^2,$$

de donde

$$\frac{M}{2m} |x_{n+1} - c| \leqslant \left(\frac{M}{2m} |x_n - c| \right)^2, n = 1, 2, 3, \ldots .$$

Aplicando sucesivamente esta desigualdad, tendremos:

$$\frac{M}{2m} |x_n - c| \leqslant \left(\frac{M}{2m} |x_{n-1} - c| \right)^2 \leqslant$$

$$\leqslant \left[\left(\frac{M}{2m} |x_{n-2} - c| \right)^2 \right]^2 \leqslant \ldots \leqslant \left(\frac{M}{2m} |b - c| \right)^{2^n}$$

Si la aproximación inicial b se elige de modo tal que sea $q \overset{\text{def}}{=} \frac{M}{2m} |b - c| < 1$, se obtendrá

$$|x_n - c| < \frac{2m}{M} q^{2^n}, n = 1, 2, \ldots ,$$

es decir, la velocidad de convergencia de las soluciones aproximadas x_n hacia la raíz $x = c$ es considerablemente superior a la de decrecimiento de la progresión geométrica cuyo denominador es, en valor absoluto, menor que la unidad.

Ejemplo. Apliquemos el método de Newton para el cálculo aproximado de una raíz de k-ésimo grado del número $a > 0$, k es entero y positivo. En este caso se trata de la solución aproximada de la ecuación $x^k - a = 0$, es decir, la fórmula (60.15) se la debe aplicar a la función $f(x) = x^k - a$.

Tenemos $f'(x) = kx^{k-1}$, y, por lo tanto, para los valores aproximados sucesivos x_n de la raíz $\sqrt[k]{x}$ existe la siguiente fórmula recurrente

$$x_{n+1} = x_n - \frac{x_n^k - a}{kx_n^{k-1}},$$

o bien

$$x_{n+1} = \frac{1}{k}\left[(k-1)x_n + \frac{a}{x_n^{k-1}}\right].$$

Con esta fórmula nos hemos encontrado para $k = 2$ en el p. 4.9.

60.3. INTERPOLACIÓN DE LAS FUNCIONES

Sea dada en el segmento $[a, b]$ una función f y sean fijos $n + 1$ valores del argumento x_i, $i = 1, 2, \ldots, n+1$:

$$a \leqslant x_1 < x_2 < \ldots < x_{n+1} \leqslant b. \tag{60.16}$$

Uno de los más sencillos problemas de interpolación consiste en la búsqueda de un polinomio $P(x)$, de grado no superior al número dado m, el cual para los valores del argumento $x = x_i$, $i = 1, 2, \ldots, n+1$, llamados *nudos de interpolación*, toma los mismos valores que la función dada f, es decir, tienen lugar las igualdades

$$f(x_i) = P(x_i), \; i = 1, 2, \ldots, n+1. \tag{60.17}$$

Tal polinomio $P(x)$ se denomina *polinomio de interpolación* que interpola la función f en los nudos dados de interpolación.

Con el objeto de investigar la cuestión de existencia del polinomio interpolador $P(x)$ que satisfaga las condiciones (60.17), escribámoslo con los coeficientes indeterminados a_j, $j = 0, 1, \ldots, m$;

$$P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m$$

y sustituyámoslo en el sistema (60.17). Obtendremos un sistema de $(n + 1)$ ecuaciones lineales con $m + 1$ incógnitas $a_0, a_1, \ldots, a_m$:

$$\begin{aligned} &a_0 + a_1x_1 + \ldots + a_mx_1^m = f(x_1) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_0 + a_1x_{n+1} + \ldots + a_mx_{n+1}^m = f(x_{n+1}). \end{aligned} \tag{60.18}$$

El *determinante* formado de los coeficientes de este sistema, ubicados en las primeras k líneas y las primeras k columnas, $k \leqslant$ mín $\{m + 1, n + 1\}$ (el número de las líneas es $n + 1$, el de las columnas es $m + 1$) lleva el nombre *de Vandermonde* y es conocido a partir del curso de álgebra:

$$W(x_1, \ldots, x_k) = \begin{vmatrix} 1 & x_1 \ldots x_1^{k-1} \\ 1 & x_2 \ldots x_2^{k-1} \\ \ldots\ldots & \ldots\ldots \\ 1 & x_k \ldots x_k^{k-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant k} (x_j - x_i).$$

En el caso dado este determinante es distinto de cero, pues todos los nudos de interpolación son diferentes. Por eso el rango de la matriz de coeficientes del sistema (60.18) es igual al mínimo de los dos números $m + 1$ y $n + 1$. Si $n > m$, el sistema (60.18), en el caso general, no tiene solución. Si $n \leqslant m$, la solución del sistema (60.18) siempre existe, con la particularidad de que, cuando $n = m$, la solución es única, y cuando $n < m$, hay una infinidad de soluciones. De este modo, *cualesquiera que sean los valores dados en $n + 1$ nudos* (60.16), *siempre existe un polinomio (y, además, el único) de grado no superior a n, que toma en los nudos mencionados los valores dados.*

Para hallar el polinomio de interpolación $P(x)$, se puede resolver el sistema (60.18). Sin embargo, puede utilizarse también otro procedimiento, más corto. Examinemos un polinomio

$$P_i(x) = \frac{(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_{n+1})}{(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_{n+1})}, \quad i = 1, 2, \dots, n + 1.$$

Es evidente que $P_i(x)$ es un polinomio de grado n y que

$$P_i(x_i) = 1, \; P_i(x_j) = 0,$$
$$i = 1, 2, \dots, n + 1, \; j = 1, 2, \dots, i - 1, i + 1, \dots, n + 1. \tag{60.19}$$

Por eso el polinomio de interpolación buscado puede ser escrito en la forma

$$P(x) = \sum_{i=1}^{n+1} f(x_i) P_i(x). \tag{60.20}$$

En efecto, la expresión escrita es un polinomio de grado no superior a n, y, en virtud de (60.19), satisface las condiciones (60.17).

El *polinomio de interpolación* escrito en la forma (60.20) lleva el nombre *de Lagrange*.

Investiguemos ahora la diferencia entre la función y el polinomio de interpolación:

$$R(x) = f(x) - P(x)$$

que se denomina *término residual de la interpolación*. Supongamos que la función f es $n + 1$ veces derivable en el segmento $[a, b]$. Entonces esta misma propiedad la posee también el resto $R(x)$, con la particularidad de que

$$R^{(n+1)}(x) = f^{(n+1)}(x), \; a \leqslant x \leqslant b, \tag{60.21}$$

pues $P^{(n+1)}x = 0$. Pongamos

$$\omega(x) = (x - x_1)(x - x_2) \dots (x - x_{n+1}),$$

fijemos $x \in [a, b]$ y analicemos una función auxiliar

$$\varphi(t) = R(t) - \frac{R(x)}{\omega(x)} \omega(t), \; a \leqslant t \leqslant b.$$

La función $\varphi(t)$, evidentemente, es también n + 1 veces derivable en el segmento $[a, b]$, además de (60.21) y de lo que $\omega^{(n+1)}(t) = (n+1)!$ tenemos

$$\varphi^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)!\frac{R(x)}{\omega(x)}. \tag{60.22}$$

Ahora, la función $\varphi(t)$ se anula en $n + 2$ puntos $x, x_1, x_2, \ldots, x_{n+1}$; por eso, en virtud del teorema de Rolle, su derivada se reduce a cero por lo menos en $n + 1$ puntos del segmento $[a, b]$, la segunda derivada en n puntos, etc. Por inducción llegamos a que la $(n + 1)$-ésima derivada de la función φ se reduce a cero por lo menos una vez dentro del segmento $[a, b]$. Sea $\varphi^{(n+1)}(\zeta) = 0$, $a < \zeta < b$, entonces de (60.22) obtendremos

$$R(x) = \frac{\omega(x)}{(n+1)!} f^{(n+1)}(\zeta),$$

o bien, más detalladamente,

$$R(x) = \frac{(x - x_1)(x - x_2)\ldots(x - x_{n+1})}{(n+1)!} f^{(n+1)}(\zeta), \quad a \leqslant x \leqslant b, \; a < \zeta < b.$$

De aquí proviene la estimación del término residual

$$|R(x)| \leqslant \frac{1}{(n+1)!} \max_{a \leqslant x \leqslant b} |(x - x_1)(x - x_2)\ldots(x - x_{n+1})| \sup_{a \leqslant x \leqslant b} |f^{(n+1)}(x)|.$$

Hemos de notar que en general, incluso para las funciones analíticas en el segmento $[a, b]$, el término residual de interpolación no tiende a cero en dicho segmento para $n \to \infty$, es decir, los polinomios de interpolación no convergen hacia la propia función. La construcción de los ejemplos correspondientes es bastante engorrosa, razón por la cual no nos detendremos en esta afirmación.

60.4. FÓRMULAS DE CUADRATURA

Consideraremos ahora algunos métodos de la integración aproximada de las funciones. Las fórmulas para los valores aproximados de las integrales se denominan *fórmulas de cuadratura*.

Supongamos que en el segmento $[a, b]$ se ha dado una función f. Dividamos el segmento $[a, b]$ en n partes iguales mediante los puntos x_k, $k = 1, 2, \ldots, n - 1$:

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b; x_k - x_{k-1} = \frac{b - a}{n}, \; k = 1, 2, \ldots, n.$$

Las fórmulas de cuadratura a considerar se obtendrán sustituyendo en el proceso de integración, la función f en cada segmento $[x_{k-1}, x_k]$ por un polinomio de interpolación de grado n. Estudiaremos los casos de $n = 0, 1, 2$. Los valores aproximados

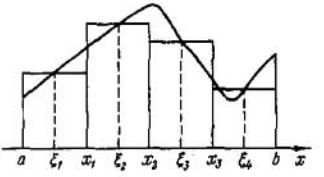


Fig. 250

correspondientes de la integral de la función f de designarán mediante el símbolo $L_n(f)$, $n = 0, 1, 2$. En el primer caso (cuando $n = 0$) la fórmula de cuadratura correspondiente se denomina *fórmula de los rectángulos*, en el segundo caso (cuando $n = 1$), *fórmula de los trapecios*, en el tercer caso (cuando $n = 2$), *fórmula de las parábolas* o, con mayor frecuencia, *fórmula de Simpson*.

FÓRMULA DE LOS RECTÁNGULOS

Para interpolar la función f en el segmento $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$, mediante un polinomio de grado nulo, es suficiente prefijar un solo nudo. Tomemos a título de nudo el centro del segmento $[x_{k-1}, x_k]$:

$$\xi_k = \frac{x_{k-1} + x_k}{2}.$$

El polinomio de interpolación será una constante

$$P_k(x) = f(\xi_k), \; k = 1, 2, \ldots, n.$$

Con tal interpolación sustituimos la función dada f por una "función escalonada", con más precisión, por un surtido de funciones, que son constantes en cada segmento $[x_{k-1}, x_k]$ e iguales al valor de la función f en el centro de dicho segmento (fig. 250). En lugar de la integral $\int_{x_{k-1}}^{x_k} f(x)dx$ tomemos otra integral $\int_{x_{k-1}}^{x_k} P_k(x)dx$, es decir, sustituyamos el área del trapecio curvilíneo por la del rectángulo correspondiente.

Escribamos ahora la fórmula de cuadratura de los rectángulos:

$$L_0[f] = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} P_k(x)dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(\xi_k)dx = \frac{b-a}{n} \sum_{k=1}^{n} f(\xi_k); \tag{60.23}$$

así pues,

$$L_0[f] = \frac{b-a}{n}\left[f\left(\frac{a+x_1}{2}\right) + f\left(\frac{x_1+x_2}{2}\right) + \ldots + f\left(\frac{x_{n-1}+b}{2}\right)\right].$$

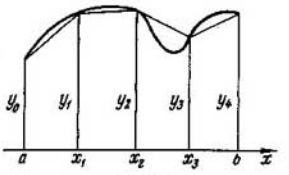


Fig. 251

FÓRMULA DE LOS TRAPECIOS

Tomemos en cada segmento $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$, un polinomio interpolador $P_k(x)$ de primer grado determinado por los nudos de interpolación x_{k-1} y x_k. Suponiendo $y_i = f(x_i)$, $i = 0, 1, \ldots, n$, obtendremos (véase (60.20))

$$P_k(x) = \frac{x - x_k}{x_{k-1} - x_k} y_{k-1} + \frac{x - x_{k-1}}{x_k - x_{k-1}} y_k, \quad k = 1, 2, \ldots, n.$$

De este modo sustituimos la función dada f por una función lineal a trozos. En lugar de la integral $\int_{x_{k-1}}^{x_k} f(x)dx$ tomemos otra integral $\int_{x_{k-1}}^{x_k} P_k(x)dx$, es decir, sustituyamos el área del trapecio curvilíneo por el área correspondiente de un trrapecio ordinario (fig. 251).

Observando que

$$\int_{x_{k-1}}^{x_k} P_k(x)dx = \frac{y_{k-1} + y_k}{2} \frac{b - a}{n}, \quad k = 1, 2, \ldots, n$$

obtendremos la fórmula de cuadratura de los trapecios

$$L_1[f] = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} P_k(x)dx = \frac{b - a}{n} \sum_{k=1}^{n} \frac{y_{k-1} + y_k}{2}, \tag{60.24}$$

o bien

$$L_1[f] = \frac{b - a}{n} \left[\frac{f(a) + f(b)}{2} + f(x_1) + f(x_2) + \ldots + f(x_n) \right].$$

FÓRMULA DE SIMPSON

Tomemos en cada segmento $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$, un polinomio de interpolación $P_k(x)$ de segundo grado determinado por los nudos de interpolación x_{k-1},

$\xi_k = \frac{x_{k-1} + x_k}{2}$ y x_k. Entonces

$$P_k(x) = \frac{(x - \xi_k)(x - x_k)}{(x_{k-1} - \xi_k)(x_{k-1} - x_k)} f(x_{k-1}) + \\ + \frac{(x - x_{k-1})(x - x_k)}{(\xi_k - x_{k-1})(\xi_k - x_k)} f(\xi_k) + \frac{(x - x_{k-1})(x - \xi_k)}{(x_k - x_{k-1})(x_k - \xi_k)} f(x_k).$$

Por el cálculo directo nos convencemos de que

$$\int_{x_{k-1}}^{x_k} \frac{(x - \xi_k)(x - x_k)}{(x_{k-1} - \xi_k)(x_{k-1} - x_k)} dx = \frac{1}{6}(x_k - x_{k-1}) = \frac{1}{6}\frac{b-a}{n},$$

$$\int_{x_{k-1}}^{x_k} \frac{(x - x_{k-1})(x - x_k)}{(\xi_k - x_{k-1})(\xi_k - x_k)} dx = \frac{2}{3}(x_k - x_{k-1}) = \frac{2}{3}\frac{b-a}{n},$$

$$\int_{x_{k-1}}^{x_k} \frac{(x - x_{k-1})(x - \xi_k)}{(x_k - x_{k-1})(x_k - \xi_k)} dx = \frac{1}{6}(x_k - x_{k-1}) = \frac{1}{6}\frac{b-a}{n},$$

por lo cual

$$\int_{x_{k-1}}^{x_k} P_k(x)dx = \frac{b-a}{n}\left[\frac{1}{6}f(x_{k-1}) + \frac{2}{3}f(\xi_k) + \frac{1}{6}f(x_k)\right].$$

Ahora ya no será difícil escribir la fórmula de cuadratura de Simpson:

$$L_2[f] = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} P_k(x)dx = \frac{b-a}{n} \sum_{k=1}^{n} \left[\frac{1}{6}f(x_{k-1}) + \frac{2}{3}f(\xi_k) + \frac{1}{6}f(x_k)\right] \quad (60.25)$$

o bien

$$L_2[f] = \frac{b-a}{6n}[f(a) + f(b) + 2[f(x_1) + \ldots + f(x_{n-1})] + \\ + 4[f(\xi_1) + \ldots + f(\xi_n)]].$$

60.5. ERROR DE LAS FÓRMULAS DE CUADRATURA*)

Hemos visto que en todos los tres casos examinados las fórmulas de cuadratura (véanse (60.23), (60.24), (60.25)) tienen por expresión

*) En este punto seguimos las ideas desarrolladas en la monografía de S. M. Nikolski "Fórmulas de cuadratura".

$$L(x) = \sum_{k=1}^{n} l_k(f), \tag{60.26}$$

donde

$$l_k(f) = \frac{b-a}{n} \sum_{i=0}^{m} p_i f(\xi_{ki}), \tag{60.27}$$

$x_{k-1} \leqslant \xi_{ki} \leqslant x_k$, $k = 1, 2, \ldots, n$, $i = 0, 1, \ldots, m$ mientras que p_i son ciertos números.

Para el caso de la fórmula de los rectángulos tuvimos

$$m = 0, p_0 = 1, \xi_{k0} = \frac{x_{k-1} + x_k}{2};$$

para el caso de la fórmula de los trapecios

$$m = 1, p_0 = p_1 = \frac{1}{2}, \xi_{k0} = x_{k-1}, \xi_{k1} = x_k;$$

y en el caso de la fórmula de Simpson

$$m = 2, p_0 = p_2 = \frac{1}{6}, p_1 = \frac{2}{3}, \xi_{k0} = x_{k-1}, \xi_{k1} = \frac{x_{k-1} + x_k}{2}, \xi_{k2} = x_k,$$
$$k = 1, 2, \ldots, n.$$

Ahora, sean dados algunos números p_i, llamados *pesos* y supongamos que en el segmento $[0, 1]$ viene dado un sistema de los puntos ξ_i, $i = 0, 1, \ldots, m$, llamados *nudos*. Supongamos también, como hasta ahora, que el segmento $[a, b]$ está dividido mediante los puntos x_k, $k = 0, 1, \ldots, n$, en n segmentos iguales $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$, y los puntos ξ_{ki} se obtienen de los nudos ξ_i en el transcurso de una aplicación lineal del segmento $[0, 1]$ sobre el segmento $[x_{k-1}, x_k]$, durante el cual el punto cero pasa al punto x_{k-1}, es decir, en el proceso de la aplicación $x = (x_k - x_{k-1})t + x_{k-1}$, $0 \leqslant t \leqslant 1$.

La expresión (60.26) se denomina en este caso fórmula de cuadratura correspondiente a los nudos ξ_i y los pesos p_i, $i = 0, 1, \ldots, m$.

Toda fórmula de cuadratura (60.26) posee la propiedad de linealidad: para cualesquiera dos funciones f y g, definidas en el segmento $[a, b]$ y para cualesquiera dos números λ y μ se verifica evidentemente, la igualdad

$$L(\lambda f + \mu g) = \lambda L(f) + \mu L(g).$$

Definición. *La fórmula* $L(f) = \sum_{k=1}^{n} l_k(f)$ *se llama exacta para los polinomios de grado* r, *si para todo polinomio* $P(x)$ *de grado no superior a* r, *para cualquier segmento* $[a, b]$ *y para todo número* n *(es decir, para cualquier partición del segmento* $[a, b]$ *en segmentos iguales) se verifica la igualdad*

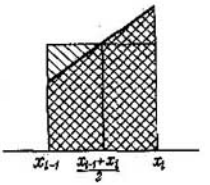


Fig. 252

$$L(P(x)) = \int_a^b P(x)\,dx.$$

Ejercicio. Demuéstrese que para que la fórmula de cuadratura $L[f]$, correspondiente a los nudos ξ_i y los pesos p_i, $i = 0, 1, \ldots, m$, sea exacta para los polinomios de grado r, es necesario y suficiente que para cualquier polinomio $P(x)$ de grado no superior a r se verifique la igualdad

$$\int_0^1 P(x)\,dx = \sum_{i=0}^{m} p_i P(\xi_i).$$

Por cuanto el polinomio de interpolación de orden r coincide, para un polinomio de grado r, con el propio polinomio, las fórmulas de cuadratura de los rectángulos, trapecios y de Simpson son exactas para los polinomios de primero, segundo y tercer grados, respectivamente.

No obstante, más aún, la fórmula de cuadratura de los rectángulos es exacta para los polinomios de primer grado, y la fórmula de Simpson, para los polinomios de tercer grado. Demostrémoslo. En efecto, en el caso de la fórmula de los rectángulos (véanse (60.23) y (60.27))

$$I_k(f) = f\left(\frac{x_{k-1} + x_k}{2}\right)(x_k - x_{k-1}).$$

El cálculo directo deja constancia de que para toda función lineal se verifica la igualdad

$$I_k(Ax + B) = \int_{x_{k-1}}^{x_k} (Ax + B)\,dx. \tag{60.28}$$

Esto se infiere claramente de la fig. 252. Al sumar las igualdades (60.28) según k desde 1 hasta n, obtendremos

$$L_0(Ax + B) = \int_a^b (Ax + B)\,dx,$$

lo que precisamente significa la exactitud de la fórmula de cuadratura de los rectángulos para los polinomios de primer grado.

En el caso de la fórmula de Simpson (véanse (60.25) y (60.27))

$$I_k(f) = \frac{b-c}{n}\left[\frac{1}{6}f(x_{k-1}) + \frac{2}{3}f\left(\frac{x_{k-1}+x_k}{2}\right) + \frac{1}{6}f(x_k)\right]. \qquad (60.29)$$

Basta mostrar que para cualquier polinomio de tercer grado $P(x)$ en este caso

$$I_k(P(x)) = \int_{x_{k-1}}^{x_k} P(x)dx, \; k = 1, 2, \ldots, n. \qquad (60.30)$$

Efectivamente, siendo demostradas estas igualdades, obtendremos, al sumarlas según k desde 1 hasta n:

$$L_2(P(x)) = \int_a^b P(x)dx,$$

es decir, la fórmula de Simpson resulta ser exacta para los polinomios de tercer grado.

Sea $P(x) = Ax^3 + Bx^2 + Cx + D$. Pongamos $Q(x) = Bx^2 + Cx + D$, entonces $P(x) = Ax^3 + Q(x)$. Por ello

$$I_k(P(x)) = AI_k(x^3) + I_k(Q(x)),$$

$$\int_{x_{k-1}}^{x_k} P(x)dx = A\int_{x_{k-1}}^{x_k} x^3dx + \int_{x_{k-1}}^{x_k} Q(x)dx, \; k = 1, 2, \ldots, n. \qquad (60.31)$$

En virtud de que la fórmula de Simpson es exacta para los polinomios de segundo grado, tenemos

$$I_k(Q(x)) = \int_{x_{k-1}}^{x_k} Q(x)dx, \; k = 1, 2, \ldots, n.$$

Por otra parte, por cálculos inmediatos nos convencemos de que

$$\int_{x_{k-1}}^{x_k} x^3dx = \frac{x_k^4 - x_{k-1}^4}{4},$$

$$I_k(x^3) = (x_k - x_{k-1})\left[\frac{x_{k-1}^3}{6} + \frac{2}{3}\left(\frac{x_{k-1}+x_k}{2}\right)^3 + \frac{x_k^3}{6}\right] = \frac{x_k^4 - x_{k-1}^4}{4}.$$

Esto demuestra precisamente la igualdad (60.30).

El orden del error de las fórmulas de cuadratura resulta relacionado con el grado de los polinomios, para de los cuales resulta exacta la fórmula de cuadratura en consideración.

Teorema. *Sea f una función r veces continuamente derivable en el segmento $[a, b]$ y sea el número $M > 0$ de tal índole que*

$$|f^{(r)}(x)| \leqslant M, a \leqslant x \leqslant b.$$

Si la fórmula de cuadratura (60.26) *es exacta para los polinomios de grado $r - 1$ ($r = 1, 2, \ldots$), entonces existe una constante $c_r > 0$, que no depende de la función f, tal que se verifica*

$$\left| \int_a^b f(x)dx - L(f) \right| \leqslant \frac{c_r M(b-a)^{r+1}}{n^r}. \tag{60.32}$$

DEMOSTRACIÓN. Representemos la función f en cada segmento $[x_{k-1}, x_k]$, de acuerdo con la fórmula de Taylor, en la forma

$$f(x) = P_k(x) + r_k(x), k = 1, 2, \ldots, n,$$

donde

$$P_k(x) = \sum_{j=0}^{r-1} \frac{f^{(j)}(x_{k-1})}{j!}(x - x_{k-1})^j$$

es un polinomio de Taylor de grado $r - 1$, y, por consiguiente, $r_k(x)$ es el término residual de la fórmula de Taylor que se escribirá en la forma de Lagrange

$$r_k(x) = \frac{f^{(r)}[x_{k-1} + \theta_k(x - x_{k-1})]}{r!}(x - x_{k-1})^r, \tag{60.33}$$

$$0 < \theta_k < 1, k = 1, 2, \ldots, n;$$

entonces

$$\int_a^b f(x)dx - L(f) = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x)dx - \sum_{k=1}^n l_k(f) =$$

$$= \sum_{k=1}^n \left[\int_{x_{k-1}}^{x_k} P_k(x)dx - l_k(P_k(x)) \right] +$$

$$+ \sum_{k=1}^n \left[\int_{x_{k-1}}^{x_k} r_k(x)dx - l_k(r_k(x)) \right]. \tag{60.34}$$

En virtud de que la fórmula de cuadratura dada es exacta para los polinomios de grado $r - 1$, se verifica la igualdad

$$\int_{x_{k-1}}^{x_k} P_k(x)dx - l_k(P_k(x)) = 0, \quad k = 1, 2, \ldots, n^*).$$

Por esto de (60.34) se infiere que

$$\left| \int_a^b f(x)dx - L(f) \right| \leqslant \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} |r_k(x)|\,dx + \sum_{k=1}^{n} |l_k(r_k(x))|. \tag{60.35}$$

Luego, de (60.33) tenemos

$$|r_k(x)| \leqslant \frac{M}{r!}\left(\frac{b-a}{n}\right)^r, \; k = 1, 2, \ldots, n.$$

Aplicando esta desigualdad, obtenemos

$$\int_{x_{k-1}}^{x_k} |r_k(x)|\,dx \leqslant \frac{M(b-a)^r}{r!n^r} \int_{x_{k-1}}^{x_k} dx = \frac{M(b-a)^{r+1}}{r!\,n^{r+1}}.$$

Suponiendo $p = \max\limits_{i=0,1,\ldots,m} |p_i|$ (véase (60.27)), tenemos

$$|l_k(r_k(x))| \leqslant \frac{b-a}{n} \sum_{i=0}^{m} |p_i|\,|r_k(\xi_{ki})| \leqslant \frac{(m+1)(b-a)^{r+1}pM}{r!\,n^{r+1}}.$$

Sustituyendo estas estimaciones en (60.35) e introduciendo las designaciones

$$c_r = \frac{1 + (m+1)p}{r!}$$

obtendremos la desigualdad (60.32). □

De la fórmula (60.32) se deduce, en particular, que al calcular las integrales con ayuda de las fórmulas de cuadratura de los rectángulos y trapecios (estas fórmulas, como se sabe, son exactas para los polinomios de primer grado, a consecuencia de lo cual podemos tomar para ellas $r = 2$) el error es de orden $O\left(\frac{1}{n^2}\right)$, y al calcular las integrales con ayuda de la fórmula de Simpson (ésta es exacta ya para los polinomios de tercer grado y podemos tomar $r = 4$) el error constituye esta vez sólo la magnitud $O\left(\frac{1}{n^4}\right)$.

*) Efectivamente, eso se deduce de la definición (véase la pág. 554) de exactitud de la fórmula de cuadratura respecto de los polinomios de grado dado, si en dicha definición a título de segmento $[a, b]$ se toma el $[x_{k-1}, x_k]$ y se pone $n = 1$.

Observemos que empleando el método citado de calcular las constantes c_r, no obtuvimos sus valores mínimos. Lo último puede lograrse perfeccionando los métodos de su cálculo.

Problema 43. Demuéstrese que para la fórmula de los rectángulos puede tomarse $c_2 = \frac{1}{24}$, para la fórmula de los trapecios $c_2 = \frac{1}{12}$, y para la fórmula de Simpson $c_4 = \frac{1}{2880}$.

60.6. CÁLCULO APROXIMADO DE LAS DERIVADAS

El cálculo aproximado de las derivadas se efectúa a base de las fórmulas que las definen. Por ejemplo, dado que

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h},$$

la así llamada razón aritmética

$$\frac{f(x + h) - f(x)}{h} \tag{60.36}$$

proporciona el valor aproximado de la derivada. Esta expresión, además, permite calcular la derivada con cualquier grado de precisión a cuenta de la elección adecuada de h, lo que se infiere de la definición de límite.

Estimemos el orden de aproximación de una derivada, calculada según la fórmula (60.36), respecto de h. Supongamos que la función f tiene en el entorno del punto x derivada acotada de segundo orden. Entonces, por la fórmula de Taylor,

$$f(x + h) = f(x) + f'(x)h + \frac{h^2}{2} f''(x + \theta h), \quad 0 < \theta < 1;$$

de aquí

$$\frac{f(x + h) - f(x)}{h} = f'(x) + \frac{h}{2} f''(x + \theta h), \quad 0 < \theta < 1,$$

es decir,

$$\frac{f(x + h) - f(x)}{h} = f'(x) + O(h), \quad h \to 0. \ \square$$

Es obvio que si en el punto x existe una derivada, entonces

$$\lim_{h \to 0} \frac{f(x + h) - f(x - h)}{2h} = f'(x).$$

Resulta que el cálculo aproximado de la derivada en un punto según la fórmula aproximada

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \tag{60.37}$$

asegura un orden más elevado de pequeñez del error respecto de h. Mostrémoslo. Supongamos que la función f tiene en el entorno del punto x la tercera derivada acotada. En este caso, según la fórmula de Taylor se tiene

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{1}{6}f'''(x+\theta_1 h)h^3,\ 0 < \theta_1 < 1,$$

$$f(x-h) = f(x) - f'(x)h + \frac{1}{2}f''(x)h^2 - \frac{1}{6}f'''(x+\theta_2 h)h^3,\ 0 < \theta_2 < 1.$$

Restando la segunda igualdad de la primera y dividiendo por $2h$, obtendremos:

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{1}{6}[f'''(x+\theta_1 h) + f'''(x+\theta_2 h)]h^2 =$$

$$= f'(x) + O(h^2),\ h \to 0.$$

De este modo, la razón aritmética (60.37) aproxima la derivada en un orden mejor que (60.36).

Para el cálculo aproximado de la segunda derivada en el punto x se puede proceder de la manera siguiente: calcular aproximadamente la primera derivada en los puntos x y $x + h$, por ejemplo, según la fórmula (60.36):

$$f'(x) \approx \frac{f(x+h) - f(x)}{h},\ f'(x+h) \approx \frac{f(x+2h) - f(x+h)}{h};$$

entonces

$$\frac{f'(x+h) - f'(x)}{h} \approx \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}.$$

La razón aritmética, que figura en el segundo miembro de la fórmula obtenida, interviene como valor aproximado de la segunda derivada en el punto x.

Cuando la función f tiene en el entorno del punto x la tercera derivada acotada, entonces al desarollar el numerador según la fórmula de Taylor, obtendremos

$$\frac{f(x+2h) - 2f(x+h) + f(x)}{h^2} = f''(x) + O(h),\ h \to 0. \tag{60.38}$$

Por analogía con el caso de la primera derivada se puede probar (en el supuesto de que la cuarta derivada en el entorno del punto x está acotada) que

$$\frac{f(x+h)-2f(x)+f(x-h)}{h^2}=f''(x)+O(h^2),\ h\to 0, \qquad (60.39)$$

es decir, el error de la fórmula aproximada (60.39) para el cálculo de la segunda derivada es en un orden mejor que el de la fórmula (60.38).

Del modo análogo se calculan las derivadas de órdenes superiores y las derivadas parciales de las funciones de varias variables.

§ 61. PARTICIÓN DEL CONJUNTO EN CLASES DE ELEMENTOS EQUIVALENTES

Varias veces en nuestro curso hemos tropezado con el concepto de equivalencia: funciones equivalentes infinitamente pequeñas e infinitamente grandes (p. 8.3), aplicaciones equivalentes de un segmento (p. 16.2) y de una región (p. 50.1), sucesiones fundamentales equivalentes de los espacios métricos (p. 57.1), funciones equivalentes en la construcción del espacio $\widetilde{RL}_2$ (p. 57.10), etc. En todos los casos mencionados la relación de equivalencia disponía de las siguientes tres propiedades: si los elementos de un conjunto en consideración se designan mediante las letras x, y, z, ... , y los elementos equivalentes x e y, mediante el símbolo $x \sim y$, entonces:

1. Todo elemento del conjuntio en consideración es equivalente a sí mismo: $x \sim x$ (*reflexi*vidad).
2. Si $x \sim y$, entonces también $y \sim x$ (simetría).
3. Si $x \sim y$ e $y \sim z$, entonces $x \sim z$ (transitividad).

Siempre se suponía sin ninguna duda que un conjunto de tales o cuales elementos en el que se ha introducido el concepto de equivalencia y que posee las propiedades de reflexividad, simetría y transitividad se descompone en las clases disjuntas de elementos equivalentes. Esto es así en realidad. Enunciemos y demostremos esta afirmación para el caso general.

Sean dados un conjunto $A = \{x, y, z, \dots\}$ y un subconjunto del conjunto de pares ordenados de sus elementos que posee las siguientes propiedades: si el par (x, y) pertenece a dicho subconjunto, los elementos x e y se llaman equivalentes y se escribe $x \sim y$, con la particularidad de que se cumplen las propiedades de reflexividad, simetría y transitividad. En tal caso suele decirse que en el conjunto A se ha dado una relación de equivalencia.

Teorema. *Si en un conjunto se ha dado una relación de equivalencia, dicho conjunto es la suma de sus subconjuntos disjuntos dos a dos de los elementos equivalentes entre sí.*

DEMOSTRACIÓN. Sea $A = \{x, y, z, \dots\}$ un conjunto en el cual se ha dado una relación de equivalencia. Para todo elemento $x \in A$ designemos mediante A_x el conjunto de todos los elementos del conjunto A equivalentes al elemento x. Mostremos que

$$A = \bigcup_{x \in A} A_x \qquad (61.1)$$

y que la representación citada del conjunto A en forma de una suma de los subconjuntos A_x es la que se busca, es decir, que los sumandos A_x son disjuntos dos a dos.

Ante todo, en virtud de que la relación de equivalencia posee la propiedad de reflexividad, para todo $x \in A$ tenemos $x \sim x$, y, por lo tanto, $x \in A_x$, es decir, todo elemento del conjunto A pertenece a cierto A_x, por lo cual

$$A \subset \bigcup_{x \in A} A_x. \tag{61.2}$$

Por otra parte, todo elemento del conjunto A_x es, debido a la propia construcción, un elemento del conjunto A. Por consiguiente, $A_x \subset A$ y por esta razón

$$\bigcup_{x \in A} A_x \subset A. \tag{61.3}$$

De las inclusiones (61.2) y (61.3) se deduce la igualdad (61.1).

Demostraremos ahora que cualesquiera dos elementos de cada uno de los conjuntos A_x son equivalentes entre sí. Efectivamente, sea $y \in A_x$, $z \in A_x$; esto significa que $y \sim x$ y $z \sim x$. Por ser simétrica la relación de equivalencia, de aquí se infiere que $x \sim z$, de donde, gracias a la transitividad, $y \sim z$.

Probemos por fin que los sumandos en el segundo miembro de la igualdad (61.1) son disjuntos dos a dos. A saber, mostremos que para cualesquiera dos elementos x' y x'' los conjuntos $A_{x'}$ y $A_{x''}$ o bien coinciden, o bien no se intersecan. En efecto, supongamos que los conjuntos $A_{x'}$ y $A_{x''}$ tengan por lo menos un elemento común: $y \in A_{x'} \cap A_{x''}$ y sea $z \in A_{x''}$. Por cuanto ya tenemos demostrado el hecho de que para todo elemento del conjunto A_x sus dos elementos cualesquiera son equivalentes, entonces $z \sim y$, $y \sim x''$, y, por ende, $z \sim x''$, es decir, $z \in A_{x''}$. El elemento z es un elemento arbitrario del conjunto $A_{x''}$, razón por la cual

$$A_{x'} \subset A_{x''}; \tag{61.4}$$

análogamente

$$A_{x''} \subset A_{x'}. \tag{61.5}$$

De (61.4) y (61.5) se sigue que

$$A_{x'} = A_{x''}.$$

De este modo, si los conjuntos $A_{x'}$ y $A_{x''}$ tienen por lo menos un elemento común, ellos coinciden; si tal elemento está ausente, los conjuntos citados son, evidentemente, disjuntos.

Así pues, la representación (61.2) posee en realidad todas las propiedades enunciadas en el teorema. □

§ 62*. LÍMITE SEGÚN UN FILTRO

Al estudiar el curso del análisis nos encontramos con dos conceptos de límite: el límite de una función (límite de una sucesión, como caso particular del límite citado) y el de las sumas integrales. Resulta que existe un concepto de límite más general, llamado límite según un filtro, el cual contiene en sí mismo ambos conceptos men-

cionados de límite como casos particulares. La existencia de tal concepto nos ofrece, por supuesto, una satisfacción estética, razón por la cual daremos a conocer en este párrafo su definición. No obstante, la introducción de este concepto no proporciona, en esencia, ventajas algunas para el estudio del análisis matemático, por lo que se explica el hecho de que dicho concepto viene al final del curso.

62.1. ESPACIOS TOPOLÓGICOS

Definición 1. *Supongamos que X es cierto conjunto y en él viene dado un sistema $\Omega = \{G\}$ de subconjuntos que satisfacen las siguientes condiciones:*

1°. La intersección de un número finito de conjuntos del sistema Ω pertenece a este sistema.

2°. La unión de cualquier totalidad de conjuntos del sistema Ω pertenece a este sistema.

3°. $X \in \Omega$, $\varnothing \in \Omega$.

En estas condiciones el conjunto X se llama espacio topológico, el sistema Ω lleva el nombre de su topología y los conjuntos del sistema Ω son sus subconjuntos abiertos.

Para cualquier punto $x \in X$, todo conjunto $G \in \Omega$ que contiene dicho punto se llama entorno de éste.

*Si cualesquiera dos puntos de un espacio topológico tienen entornos que no se intersecan, el espacio lleva el nombre de Hausdorff**).

A título de espacio topológico de Hausdorff interviene todo espacio métrico, puesto que sus subconjuntos abiertos forman un sistema que satisface las condiciones 1°, 2°, 3° de la definición 1 (véase el p. 57.1). Existen también los así llamados espacios topológicos no metrizables (véase "Introducción a la teoría de los conjuntos y la topología general" de P.S. Alexándrov, M., 1977).

Para cualquier punto $x \in X$, todo entorno de x no es, a ciencia cierta, un conjunto vacío, puesto que contiene por lo menos un elemento, a saber, el propio punto x.

Definición 2. *Todo subsistema $\mathfrak{B}$ del sistema Ω de conjuntos abiertos de un espacio topológico recibe el nombre de base de la topología del espacio citado, siempre que cualquier conjunto abierto no vacío del espacio (es decir, un conjunto no vacío del sistema Ω) es una unión de cierta totalidad de conjuntos de $\mathfrak{B}$.*

Así por ejemplo, en un espacio métrico la base de la topología la constituye el conjunto $\mathfrak{B}$ de todos los ε-entornos de todos los puntos de este espacio. Efectivamente, cualquiera que sea el conjunto abierto no vacío G del espacio métrico dado, para todo su punto $x \in G$ existe tal $\varepsilon > 0$ que el ε-entorno de este punto está contenido en G: $U(x, \varepsilon) \subset G$. Elijamos y fijemos para todo punto $x \in G$ uno de tales entornos, entonces el conjunto G será, evidentemente, su unión:

$$G = \bigcup_{x \in G} U(x, \varepsilon).$$

*) F. Hausdorff (1868 — 1942), matemático alemán.

Ejercicio 1. Demuéstrese que en cualquier espacio métrico el conjunto de todos los ε-entornos (con ε racional) de todos los puntos de este espacio forma su base de topología.

La topología puede ser definida con ayuda de la base de topología. A saber, si $\mathfrak{B} = \{A\}$ es la base de topología Ω del espacio X, entonces, de conformidad con la definición 2, Ω es el sistema de todos los subconjuntos del espacio X, cada uno de los cuales o bien es una unión de cierta totalidad de conjuntos de $\mathfrak{B}$, o bien es vacío.

Definición 3. *El sistema $\mathfrak{B}(x)$ de entornos del punto x de un espacio topológico X se llama base local de la topología en dicho punto, si, cualquiera que sea el entorno V del punto x en el espacio X, existe tal entorno $U \in \mathfrak{B}(x)$ que*

$$U \subset V.$$

Es obvio que la totalidad de todos los entornos del punto dado forma su base local de la topología. Para cualquier punto de un espacio métrico su base local de la topología la forman también, por ejemplo, todos sus ε-entornos de radio $\varepsilon = \frac{1}{n}$, $n = 1, 2, \ldots$.

La unión de bases locales de la topología en todos los puntos forma la base de la topología de todo el espacio, pues cualquier conjunto abierto no vacío puede representarse como una unión de los entornos (pertenecientes al conjunto) de sus puntos, donde los entornos mencionados se toman de las bases locales de la topología en consideración. De este modo, la topología en un conjunto puede prefijarse definiendo las bases locales de la topología en cada uno de los puntos del conjunto.

Con ayuda del concepto de entorno para los espacios topológicos se introducen (textualmente, al igual que para los espacios métricos, véanse el p. 57.1 y p. 18.2) nociones de puntos adherente, límite y aislado, como también la de conjunto cerrado.

62.2. FILTROS

En lo sucesivo mediante $\mathfrak{B}(X)$ se designará el conjunto de todos los subconjuntos del conjunto X.

Definición 4. *Sea X un conjunto no vacío. El conjunto $\mathfrak{F} \subset \mathfrak{B}(X)$ recibe el nombre de filtro (o, más detalladamente, filtro en el conjunto X), siempre que:*

1° *Para cualesquiera $A' \in \mathfrak{F}$ y $A'' \in \mathfrak{F}$ existe tal $A \in \mathfrak{F}$ que*

$$A \subset A' \cap A''.$$

2°. $\varnothing \notin \mathfrak{F}$, $\mathfrak{F} \neq \varnothing$.

De las propiedades 1° y 2° se desprende que la intersección de cualquier número finito de conjuntos pertenecientes al filtro no es vacío.

Ejemplos. 1. Sea $X \neq \varnothing$, $X \supset A_0 \neq \varnothing$. Entonces, el conjunto $\mathfrak{F} = \{A : A_0 \subset A \in \mathfrak{B}(X)\}$ es un filtro en X. Efectivamente, es evidente que $A_0 \neq \mathfrak{F}$, y si $A' \in \mathfrak{F}$ y $A'' \in \mathfrak{F}$, se tiene $A' \cap A'' \supset A_0 \neq \varnothing$, es decir, ambas condiciones 1° y 2° de la definición 4 quedan cumplidas.

2. Sea $x \in X$. El conjunto $\mathfrak{F} = \{A : x \in A \in \mathfrak{B}(X)\}$ es un filtro en X. Este filtro es un caso particular del filtro considerado en el ejemplo antecedente en que el conjunto A_0 se componía de un solo punto x.

3. Sea $X = N$ un conjunto de todos los números naturales y

$$A_n = \{m : m \in N, m > n\}, n \in N. \tag{62.1}$$

Entonces el conjunto A_n forma un filtro, designado por $F_N = \{A_n\}$ y llamado *filtro natural*.

Comprobemos que F_N es un filtro. En efecto, $N \in F_N$ y, por consiguiente $F_N \neq \varnothing$, todos los $A_n \neq \varnothing$, y si $m < n$, se tiene $A_n \cap A_m = A_n \in F_N$.

4. Sea de nuevo $X = N$. El sistema de subconjuntos $\mathfrak{F}_N$ del conjunto N, cada uno de los cuales es un complemento al subconjunto finito del conjunto N, forma también un filtro en N, llamado *filtro de Fréchet* y continente en sí el filtro natural F_N.

Mostremos que $\mathfrak{F}_N$ es, de hecho, un filtro. Si $A \in \mathfrak{F}_N$, y $B \in \mathfrak{F}_N$, designemos mediante $n \in N$ el mayor de los números que integran el conjunto $(N \setminus A) \cup (N \setminus B)$. Tal número existe, puesto que el conjunto indicado se compone, en virtud de la definición de $\mathfrak{F}_N$, sólo de un conjunto finito de números. Entonces el conjunto $A_n \in F_N$ (véase (62.1)) está contenido en $A \cap B$. Luego, por cuanto el conjunto de números naturales N es numerable, mientras que $N \setminus A$, donde $A \in \mathfrak{F}_N$ es, por definición del conjunto $\mathfrak{F}_N$, finito, entonces $A \neq \varnothing$. Por fin, $N \in \mathfrak{F}_N$ y, por consiguiente, $\mathfrak{F}_N \neq \varnothing$. De este modo, $\mathfrak{F}_N$ es un filtro.

5. Sea X un espacio topológico y $x \in X$. La base local de la topología $\mathfrak{B}(x)$ forma un filtro. En efecto, ante todo, es evidente que para cualquier entorno $U \in \mathfrak{B}(x)$ se tiene $x \in U$, y por esta razón $U \neq \varnothing$. Luego, para cualesquiera $U \in \mathfrak{B}(x)$ y $V \in \mathfrak{B}(x)$ la intersección $U \cap V$ es un conjunto abierto en el que se contiene el punto x, por lo cual, de conformidad con la definición de la base local de una topología, existe tal entorno $W \in \mathfrak{B}(x)$ que $W \subset U \cap V$.

6. Supongamos que X es un espacio topológico, x es el punto límite del espacio X, $\mathfrak{B}(x)$, la base local de la topología en este punto y $\mathring{\mathfrak{B}}(x)$, un conjunto de todos los "entornos reducidos" de dicha base local de la topología, es decir, $\mathring{\mathfrak{B}}(x)$ se compone de los conjuntos

$$\mathring{U}(x) \stackrel{\text{def}}{=} U(x) \setminus \{x\}, \ U(x) \in \mathfrak{B}(x).$$

Bajo estas condiciones $\mathring{\mathfrak{B}}(x)$ forma un filtro. Efectivamente, si $\mathring{U} \in \mathring{\mathfrak{B}}(x)$, entonces por cuanto el punto x es límite para el espacio X, existe un punto $y \in \mathring{U}$, y, por ende, $\mathring{U} \neq \varnothing$. Luego, para cualesquiera $\mathring{U} \in \mathring{\mathfrak{B}}(x)$ y $\mathring{V} \in \mathring{\mathfrak{B}}(x)$ tenemos, de acuerdo con su definición: $\mathring{U} = U \setminus \{x\}$, $\mathring{V} = V \setminus \{x\}$, $U \in \mathfrak{B}(x)$, $V \in \mathfrak{B}(x)$. La intersección $U \cap V$ es un entorno del punto x, por lo cual existe tal entorno $W \in \mathfrak{B}(x)$ que $W \subset U \cap V$ y por esta razón $\mathring{W} = W \setminus \{x\} \subset \mathring{U} \cap \mathring{V}$. Así pues, $\mathring{\mathfrak{B}}(x)$ es, de hecho, un filtro.

Definición 5. *El filtro $\mathfrak{F}_1 = \{A\}$ en el conjunto X lleva el nombre de filtro más fuerte que el filtro $\mathfrak{F}_2 = \{B\}$ en el mismo conjunto, si para cualquier conjunto $B \in \mathfrak{F}_2$ existe tal $A \in \mathfrak{F}_1$ que $A \subset B$.*

Definición 6. *Si el filtro $\mathfrak{F}_1$ es más fuerte que el filtro $\mathfrak{F}_2$, y $\mathfrak{F}_2$ es más fuerte que $\mathfrak{F}_1$, los filtros $\mathfrak{F}_1$ y $\mathfrak{F}_2$ se denominan equivalents.*

Ejemplo 7. Sea $\mathfrak{B}(x)$ una base local de la topología del punto x de un espacio metrico, compuesta de todos sus ε-entornos, y sea $\mathfrak{B}_0(x)$ su base local de la topolo-

gía que contiene sólo los entornos de radio $\varepsilon = \frac{1}{n}$, $n = 1, 2, \ldots$. Los filtros $\mathfrak{B}(x)$ y $\mathfrak{B}_0(x)$ son equivalentes.

Ejercicio 2. Demuéstrese que los filtros en los ejemplos 3 y 4 son equivalentes.

Definición 7. *El filtro $\mathfrak{F}_1$ se llama subfiltro del filtro $\mathfrak{F}_2$, si todo elemento del filtro $\mathfrak{F}_1$ es también un elemento del filtro $\mathfrak{F}_2$, es decir, si $\mathfrak{F}_1 \subset \mathfrak{F}_2$.*

Es obvio que un filtro es más fuerte que cualquiera de sus subfiltros.

Definición 8. *Todo subfiltro de un filtro, equivalente al propio filtro, se denomina base del filtro.*

En el ejemplo 7 el filtro $\mathfrak{B}_0(x)$ es la base del filtro $\mathfrak{B}(x)$, mientras que el filtro natural F_N es la base del filtro de Fréchnet $\mathfrak{F}_N$, construido en el ejemplo 4.

A veces resulta cómodo considerar los filtros que satisfacen una condición adicional más.

Definición 9. *El filtro $\mathfrak{F}$ en un conjunto X se denomina completo, si de las condiciones*

$$A \in \mathfrak{F},\ B \in \mathfrak{B}(X) \text{ y } A \subset B$$

se deduce que

$$B \in \mathfrak{F}.$$

En los ejemplos 1, 2 y 4, considerados más arriba, los filtros eran completos. En el ejemplo 4 (filtro de Fréchnet) esto se deduce de lo que si $A \in \mathfrak{F}_N$ y, por consiguiente, su complemento en el conjunto de números naturales N es finito, entonces cualquier subconjunto de números naturales B, contenido en A, tiene también un complemento finito en N, pues, si $A \subset B \subset N$, entonces $N \setminus B \subset N \setminus A$.

Entre tanto, los filtros, considerados en los ejemplos 3, 5 y 6, ya no son completos. En el ejemplo 3 el filtro natural F_N no es completo, puesto que no todo subconjunto A (que contiene un conjunto del tipo A_n, véase (62.1)) del conjunto de números naturales es del mismo tipo, es decir, pertenece al filtro natural F_N. Los filtros, considerados en los ejemplos 5 y 6, no son completos, puesto que no todo conjunto, que contiene un conjunto abierto, es por sí mismo necesariamente abierto.

A veces en la literatura matemática el filtro completo se llama sencillamente filtro, mientras que el filtro en el sentido de la definición 4, base del filtro. Esto se justifica por la validez de la siguiente afirmación.

Lema 1. *Todo filtro es una base de cierto filtro completo.*

DEMOSTRACIÓN. Sea $\mathfrak{F} = \{A\}$ un filtro en el conjunto X. Definamos el conjunto $\mathfrak{G}$ como un conjunto de todos los subconjuntos B del conjunto X de tal género que cada uno de los subconjuntos citados tenga, a título de su subconjunto, un cierto elemento del filtro $\mathfrak{F}$. Más brevemente, $B \in \mathfrak{G}$ cuando, y sólo cuando, existe tal $A \subset \mathfrak{F}$ que $A \subset B$. Mostremos que $\mathfrak{G}$ es un filtro completo y el filtro $\mathfrak{F}$ es la base de $\mathfrak{G}$.

Si $B' \in \mathfrak{G}$, $B'' \in \mathfrak{G}$, entonces existen tales $A' \in \mathfrak{F}$ y $A'' \in \mathfrak{F}$ que $A' \subset B'$, $A'' \subset B''$. Por cuanto $\mathfrak{F}$ es un filtro, se encontrará tal $A \in \mathfrak{F}$ que $A \in A' \cap A''$. Al notar que $A' \cap A'' \subset B'' \cap B'$, obtendremos $A \subset B' \cap B''$, y, por lo tanto, de acuerdo con la definición de $\mathfrak{G}$, el conjunto $B' \cap B''$ será su elemento: $B' \cap B'' \in \mathfrak{G}$. De este modo se cumple la condición 1° de la definición 4.

Si fuera $\varnothing \in \mathcal{G}$, entonces nuevamente, por definición de $\mathcal{G}$, se encontraría tal $A \in \mathcal{F}$ que $A \subset \varnothing$, pero en tal caso $A \neq \varnothing$, es decir, el conjunto vacío resultaría ser un elemento de $\mathcal{F}$ que contradiría a que $\mathcal{F}$ es un filtro. Por consiguiente, $\varnothing \in \mathcal{G}$. Además, puesto que $A \subset A$, todo conjunto $A \in \mathcal{F}$ es también un elemento de $\mathcal{G}$, es decir, $\mathcal{F} \subset \mathcal{G}$, y por cuanto $\mathcal{F}$, como cualquier otro filtro, no es vacío: $\mathcal{F} \neq \varnothing$, tampoco será vacío el conjunto $\mathcal{G}$: $\mathcal{G} \neq \varnothing$. De este modo, $\mathcal{G}$ satisface todas las condiciones de la definición 4, es decir, es un filtro. Su completitud también se deduce inmediatamente de la definición de este filtro. En efecto, si $B \in \mathcal{G}$, entonces existe tal $A \in \mathcal{F}$, que $A \subset B$. Por eso, para cada conjunto B' tal que $B \subset B' \subset X$ será válida tambien la inclusión $A \subset B'$ la que significa precisamente que $B' \in \mathcal{G}$.

Por fin, $\mathcal{F}$ es la base del filtro completo $\mathcal{G}$. Efectivamente, por un lado, como ya se ha señalado, $\mathcal{F} \subset \mathcal{G}$, es decir, el filtro $\mathcal{F}$ es un subfiltro de $\mathcal{G}$; mientras tanto, se ha notado anteriormente que todo filtro es más fuerte que cualquiera de sus subfiltros. Por otra parte, la definición del filtro $\mathcal{G}$ significa precisamente que el filtro $\mathcal{F}$ es más fuerte que el $\mathcal{G}$: cualquiera que sea $B \in \mathcal{G}$, existe un $A \in \mathcal{F}$ tal que $A \subset B$ (véase la definición 5). Así pues, los filtros $\mathcal{F}$ y $\mathcal{G}$ son equivalentes. □

Lema 2. *Sea $\mathcal{F}_1$ un filtro en el conjunto X_1 y sea $\mathcal{F}_2$ un filtro en el conjunto X_2 y*

$$\mathcal{F} \overset{\text{def}}{=} \{C : C = A \times B, A \in \mathcal{F}_1, B \in \mathcal{F}_2\}; \tag{62.2}$$

en este caso $\mathcal{F}$ es un filtro sobre el producto $X_1 \times X_2$ de los conjuntos X_1 y X_2.

El filtro $\mathcal{F}$ definido por la igualdad (62.2) se llama producto de los filtros $\mathcal{F}_1$ y $\mathcal{F}_2$. Si $\mathcal{F}$ es un producto de los filtros $\mathcal{F}_1$ y $\mathcal{F}_2$, se escribe $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$.

DEMOSTRACIÓN. Sea $C_1 \in \mathcal{F}_1$ y $C_2 \in \mathcal{F}_2$, entonces, de acuerdo con la definición (62.3), existen tales $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_1$ y $B_1 \in \mathcal{F}_2$, $B_2 \in \mathcal{F}_2$, que $C_1 = A_1 \times B_1$ y $C_2 = A_2 \times B_2$. Por cuanto $\mathcal{F}_1$ y $\mathcal{F}_2$ son filtros, se encontrarán tales $A \in \mathcal{F}_1$ y $B \in \mathcal{F}_2$ que

$$A \subset A_1 \cap A_2, B \subset B_1 \cap {}_2. \tag{62.3}$$

En virtud de la misma definición (62.2) : $A \times B \in \mathcal{F}$, además de (62.3) se deduce que

$$A \times B \subset (A_1 \times B_1) \cap (A_2 \times B_2),$$

pues, si $(x, y) \in A \times B$, se tiene $x \in A$, $y \in B$. Por consiguiente, en virtud de (62.3), $x \in A_1 \cap A_2$, $y \in B_1 \in \cap B_2$, por lo cual $(x, y) \in A_1 \times B_1$ y $(x, y) \in A_2 \times B_2$, es decir,

$$(x, y) \in (A_1 \times B_1) \cap (A_2 \times B_2).$$

Por fin, todo $C = A \times B \neq \varnothing$, $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$, pues, de conformidad con la definición del filtro, $A \neq \varnothing$, $B \neq \varnothing$. De lo que $\mathcal{F}_1 \neq \varnothing$ y $\mathcal{F}_2 \neq \varnothing$, se infiere que también $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2 \neq \varnothing$.

De este modo, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ satisface la definición del filtro. □

Lema 3. *Sean X e Y unos conjuntos, $f : X \to Y$, la aplicación de X en Y y $\mathcal{F} = \{A\}$, un filtro en el conjunto X. Entonces la totalidad de todas las imágenes $f(A)$ de los conjuntos, pertenecientes a $\mathcal{F}$, es un filtro en el conjunto Y.*

El filtro $\{f(A)\}$, $A \in \mathfrak{F}$, se denomina imagen del filtro $\mathfrak{F}$ en la aplicación f y se designa mediante

$$f(\mathfrak{F}) = \{f(A)\},\ A \in \mathfrak{F}. \tag{62.4}$$

Demostremos que $f(\mathfrak{F})$ es realmente un filtro. Sea $f(A) \in f(\mathfrak{F})$, $f(B) \in f(\mathfrak{F})$, $A \in \mathfrak{F}$, $B \in \mathfrak{F}$. En este caso existe tal elemento C del filtro $\mathfrak{F}$: $C \in \mathfrak{F}$, que $C \subset A \cap B$. Por cuanto $f(C) \subset f(A \cap B) \subset f(A) \cap f(B)$, y, además, por definición del sistema $f(\mathfrak{F})$ tenemos $f(C) \in f(\mathfrak{F})$, entonces la primera condición de la definición de filtro (véase la definición 4) queda cumplida. La segunda condición está también cumplida, dado que $f(\mathfrak{F})$ se compone sólo de los elementos del tipo $f(A)$, donde $A \in \mathfrak{F}$. Por consiguiente, $f(A) \neq \varnothing$, puesto que $A \neq \varnothing$. Por fin, de lo que $\mathfrak{F} \neq \varnothing$, se deduce que también $f(\mathfrak{F}) \neq \varnothing$. □

62.3. LÍMITE DE UN FILTRO

Definición 10. *Sea X un espacio topológico, $x \in X$ y sea $\mathfrak{F}$ un filtro en X. El punto x se llama límite del filtro $\mathfrak{F}$ o su punto límite, si el filtro $\mathfrak{F}$ es más fuerte que $\mathfrak{B}(x)$ que constituye una base local de la topología en este punto.*

Si el punto x es límite del filtro $\mathfrak{F}$, se escribe

$$x = \lim \mathfrak{F}.$$

Ejemplos. Sea $X = N$ un conjunto de todos los números naturales que se considera, como siempre, con una topología discreta: todo punto $n \in N$ se entiende como conjunto abierto (en otras palabras, cualquier punto es aislado), entonces el filtro natural F_N (véase el ejemplo 3 en el p. 62.2) no tiene límite en N.

Efectivamente, ningún número $n \in N$ es un límite del filtro F_N, pues para cualquier número $n_0 \in N$ existe una base local de la topología compuesta sólo de este número n_0 y no existe $A \in F_N$, contenido en el conjunto de un solo punto $\{n_0\}$, por cuanto todo $A \in F_N$ contiene una infinidad de elementos. De este modo, el filtro F_N no es más fuerte que la base local de la topología de cualquier número $n_0 \in N$.

2. Sea $X = N \cup \{+\infty\}$, es decir, el conjunto X se ha obtenido por adición de un "punto infinito" $+\infty$ al conjunto de los números naturales N, con la particularidad de que la base local de la topología $\mathfrak{B}(+\infty)$ se compone de toda una serie de conjuntos A_n (véase (61.1)), mientras que las bases locales $\mathfrak{B}(n)$, $n \in N$, están compuestas (como hasta ahora) por el único punto n. La base de la topología en X se define como una unión de bases locales de todos sus puntos.

En el espacio $N \cup \{+\infty\}$ el filtro natural F_N tiene por límite $+\infty$. En efecto, para cualquier entorno $A_n \in \mathfrak{B}(+\infty)$ a título de elemento $A \in F_N$ tal que $A \subset A_n$ (véase la definición 10) puede tomarse el propio elemento A_n, pues $A_n \subset F_N$.

Problema 44. Demuéstrese que para que todo filtro de un espacio topológico tenga no más de un límite, es necesario y suficiente que el espacio sea de Hausdorff.

Teorema 1. *Para que el punto x sea un límite del filtro $\mathfrak{F}$ del espacio topológico X, es necesario que este punto sea un límite de cada su base y suficiente que dicho punto sea un límite por lo menos de una base del filtro.*

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que un subfiltro $\mathfrak{F}_0$ es la base del filtro $\mathfrak{F}$ del espacio X y

$$x = \lim \mathfrak{F},$$

es decir, el filtro $\mathfrak{F}$ es más fuerte que la base local de la topología $\mathfrak{B}(x)$ en el punto x. Esto significa que para cualquier entorno $U \in \mathfrak{B}(x)$ existe tal $A \in \mathfrak{F}$ que $A \subset U$. Por cuanto $\mathfrak{F}_0$ es una base del filtro $\mathfrak{F}$, para $A \in \mathfrak{F}$ citado se encontrará tal $B \in \mathfrak{F}_0$ que $B \subset 2A$, y, por consiguiente, $B \subset U$, es decir, el cubfiltro $\mathfrak{F}_0$ es también más fuerte que la base local de la topología $\mathfrak{B}(x)$, por lo cual $x = \lim \mathfrak{F}_0$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Supongamos que el subfiltro $\mathfrak{F}_0$ del filtro $\mathfrak{F}$ es una base del filtro y $x = \lim \mathfrak{F}_0$, a sea $\mathfrak{F}_0$ es más fuerte que la base local de la topología $\mathfrak{B}(x)$, entonces, el propio filtro $\mathfrak{F}$ es con mayor razón más fuerte que $\mathfrak{B}(x)$, pues cada elemento de un subfiltro es a la vez, un elemento del filtro. Por consiguiente, $x = \lim \mathfrak{F}$. □

62.4. LÍMITE DE LA APLICACIÓN SEGÚN UN FILTRO

El concepto general de límite se da por la siguiente definición.

Definición 11. *Sean X un cierto conjunto, Y un espacio topológico, $f : X \to Y$ la aplicación de X en Y, f un filtro en X.*

El punto $b \in Y$ se llama límite de la aplicación f según el filtro $\mathfrak{F}$ y se escribe

$$\lim_{\mathfrak{F}} f(x) = b,$$

siempre que el filtro $f(\mathfrak{F})$ tiene como su límite en el espacio Y el punto b.

De este modo,

$$\lim_{\mathfrak{F}} f(x) \stackrel{\text{def}}{=} \lim f(\mathfrak{F}). \tag{62.5}$$

Ejemplos. 1. Sea $X = N$ un conjunto de todos números reales, Y un espacio topológico, $f : N \to Y$, $y_n \stackrel{\text{def}}{=} f(n)$, $n \in N$, y supongamos que F_N es el filtro natural construido en el ejemplo 3, del p. 62.2, es decir, F_N se compone de los conjuntos (62.1). En este caso el límite de la aplicación f según el filtro F_N coincide con el límite ordinario de la sucesión $\{y_n\}$ en Y. En efecto, la condición $\lim_{F_N} f(n) = b$ es equivalente, de acuerdo con (62.5), a la condición $\lim f(F_N) = b$, donde $f(F_N) = \{f(A_n)\}$, $f(A_n) = \{y_m : m > n\}$. El hecho de que el límite del filtro $f(F_N)$ es igual al punto b es indicio de que para cualquier entorno $U \in \mathfrak{B}(b)$, donde $\mathfrak{B}(b)$ es una base local de la topología en el punto b, existe un elemento $f(A_{n_0})$, contenido en U, del filtro $f(F_N)$: $f(A_{n_0}) \subset U$. Por cuanto para $n > n_0$ se cumple la inclusión $n \in A_{n_0}$, y, por ende, también la inclusión $y_n = f(n) \in f(A_{n_0})$, entonces para $n > n_0$ tiene lugar la inclusión $y_n \in U$. Esto significa precisamente que $\lim\limits_{n \to \infty} y_n = b$.

2. Sean $X = N \times N$, F_N un filtro natural, $\mathfrak{F} = F_N \times F_N$ (véase (62.2)), Y un espacio topológico, $f : N \times N \to Y$, $y_{mn} \stackrel{\text{def}}{=} f(m, n)$, $m \in N$, $n \in N$; en este caso el límite $\lim_{\mathfrak{F}} (m, n)$ coincide con el límite ordinario de la sucesión doble $\{y_{mn}\}$: el punto b se llama límite $\lim\limits_{(m, n) \to \infty} y_{mn}$ de la sucesión $\{y_{mn}\}$, si para cualquier entorno U del punto b existen tales m_0 y n_0 que para $m > n_0$ y $n > n_0$ se cumple la inclusión $y_{mn} \in U$. De este modo,

$$\lim_{\mathfrak{F}} f(m, n) = \lim_{(m, n) \to \infty} y_{mn}.$$

3. Supongamos que E es un conjunto medible según Jordan en R^n, τ es una partición de este conjunto: $\tau = \{E_i\}_{i=1}^{i=k}$, $\xi_i \in E_i$, $i = 1, 2, \ldots, k$. Supongamos, ade-

más, que los elementos del conjunto X los constituyen, a su vez, toda clase de conjuntos del tipo

$$x = \{\tau, \xi_1, \ldots, \xi_k\}. \tag{62.6}$$

Para todo $\eta > 0$ designemos mediante A_η un subconjunto del conjunto X que se compone de todos aquellos elementos x, para los cuales las finuras δ_τ de las particiones τ que las integran son inferiores a η, es decir, $\delta_\tau < \eta$.

El sistema $\mathfrak{F} = \{A_\eta\}$ es un filtro en X.

Toda función real $f: E \to R$ engendra una aplicación $\varphi_f: X \to R$ según la fórmula

$$\varphi_f(x) \stackrel{\text{def}}{=} \sum_{i=1}^{k} f(\xi_i)\mu E_i,\ x = \{\tau, \xi_1, \ldots, \xi_k\}.$$

De este modo, $\varphi_f(x)$ es un valor de la correspondiente suma integral de Riemann de la función f.

El límite de la aplicación $\varphi_f: X \to R$ según el filtro $\mathfrak{F} = \{A_\eta\}$ coincide con el límite ordinario de las sumas integrales de Riemann de la función f, a condición de que las finuras de las particiones en consideración tienden a cero:

$$\lim_{\mathfrak{F}} \varphi_f(x) = \lim_{\delta_\tau \to 0} \sum_{i=1}^{k} f(\xi_i)\mu E_i.$$

4. Supongamos que X e Y son unos espacios topológicos, $f: X \to Y$, $a \in X$ y $\mathfrak{F}$ es un filtro de tal género en X que $\lim \mathfrak{F} = a$ (es decir, el filtro $\mathfrak{F}$ es más fuerte que cierta base local de la topología $\mathfrak{B}(a)$ en el punto a).

El límite $\lim_{\mathfrak{F}}(x)$ en este caso se denomina *límite de la aplicacion f según el filtro* $\mathfrak{F}$ *en el punto a*.

Elegidos adecuadamente los filtros $\mathfrak{F}$, se obtendrán, en particular, los límites en un punto dado según diferentes conjuntos. Por ejemplo, si el filtro $\mathfrak{F}$ se compone de los entornos de cierta base local de la topología $\mathfrak{B}(a)$ del punto a, entonces la existencia del límite $\lim_{\mathfrak{F}} f(x)$ en el punto a según tal filtro significa la continuidad de la aplicación f en el punto a, con la particularidad de que $\lim_{\mathfrak{F}} f(x) = \lim_{x \to a} f(x) = f(a)$.

Si el punto a es un punto límite del conjunto X, mientras que el filtro $\mathfrak{F}$ se compone de los entornos reducidos de cierta base local de la topología en este punto (véase el ejemplo 6 en el p. 62.2), entonces $\lim_{\mathfrak{F}} f(x)$ coincide con el límite habitual $\lim_{x \to a} f(x)$.

Ha de notarse que antes el símbolo $x \to a$ no tenía para nosotros un sentido individual: toda la designación $\lim_{x \to a} f(x)$ se consideraba en total. Ahora, al final del curso, vemos que el símbolo $x \to a$ puede considerarse como designación del filtro $\mathring{\mathfrak{B}}(a)$ o del filtro $\mathfrak{B}(a)$, según el cual se toma el límite de una aplicación (en el primer caso se obtendrá la definición ordinaria del límite de una aplicación en el punto a, en el segundo caso, la definición de su contunuidad en dicho punto).

Así pues, todos los conceptos de límite con los que nos hemos enfrentado anteriormente son, de hecho, los casos particulares del límite de una aplicación según un filtro.

Indice de nombres

Banach S. 441, 477
Bernoulli J. 343
Bessel F. 381, 382, 481
Buniakovski V. Ya. 448

Cantor G. 56
Cauchy A. L. 53, 66, 311, 383, 446, 447, 450, 482, 490
Cramer G. 51

Darboux G. 144, 152
Dini U. 362
Dirac P. A. 506, 510
Dirichlet L. P. G. 167, 328, 355, 356, 357, 358

Euler L. 325, 328

Fejer L. 371, 374
Fourier J. B. 346—376, 380—394, 400—412, 477—497, 526, 527—535
Fréchet M. R. 511

Gauss C. F. 285
Gram J. P. 469
Green G. 203, 206, 207, 223
Guldin P. 236

Hardy G. H. 173
Hausdorff F. 563
Heaviside O. 509, 522
Heine H. E. 53
Helmholtz H. 300
Hilbert D. 452, 477, 490
Hölder O. 368—370, 398, 435, 490

Jakobi K. G. J. 43, 71, 72, 91
Jordan C. 118, 130, 223, 225, 414

Kronecker L. 424

Lagrange J. L. 11, 19, 94, 101, 304, 331, 537, 549, 557
Laplace P. S. 87, 223, 403, 404
Lebesgue H. L. 360, 434, 466
Legendre A. M. 470—472, 476, 486
Leibniz v. G. W. 207, 304, 322, 509
Lipschitz R. 369
Littlewood J. E. 173
L'Hospital G. de 332, 339

Minkowski H. 172, 427, 434
Möbius A. F. 262, 266

Newton I. 207, 322, 345, 509, 542, 545
Nikolski S. M. 466, 553

Ostrogradski M. V. 285

Parseval M. A. 383, 483, 484, 492, 493
Peano J. G. 14, 133
Pitágoras 483
Plancherel M. 502
Poisson S. D. 227
Polya G. 173

Riemann B. 135, 223, 331, 434
Rolle M. 81, 331, 549

Schwarz K. H. 66, 386, 446, 447, 450, 490
Schwarz L. 511
Simpson T. H. 551, 552, 556, 559
Sóbolev S. L. 511
Sojotski Yu. V. 520
Stirling J. 338, 343
Stokes G. G. 290
Sylvester J. J. 32

Taylor B. 11, 12, 537—540, 557

Vandermonde A. 548

Weierstrass K. 29, 56, 310, 376, 398, 429, 476

Índice alfabético de materias

Aditividad de la medida 126
Aplicación 52
— continua 53, 54, 58, 514
— derivable 67, 74
— inversa 58
— isomorfa 425, 438, 487
— lineal 61
— localmente homeomorfa 76
— natural 460
— regular 242
— uniformemente continua 56
— vectorial 60
Aproximación 478, 480
Área (medida) de la superficie 255
Axiomas de distancia 411
— — Fréchet 511

Base de la topología 563, 564
— del espacio 424, 427
Borde de la superficie 237

Campo escalar 276
— vectorial 276
— — potencial 279, 297, 300
— — solenoidal 295, 300
Cápsula lineal del conjunto 423
Cinta (banda) de Möbius 262, 263
Circulación 278, 281, 290
Coeficientes de Fourier 348, 364, 380, 392, 479, 480
— del desarrollo del elemento 445
Complementos ortogonales 108
Completación del espacio 417, 453, 465
Condición clásica de Hölder 369
— de Lipschitz 369
Conjunto acotado 437
— cuadrable 119
— cubicable 119
— denso en el espacio 416, 443, 465
— medible según Jordan 118
Constante de encaje 475
Contorno de frontera 205
— exterior 205
— interior 206
— que limita la superficie 290
— suave a trozos 210
Convergencia en media 436
Convolución de las funciones 408, 410
Coordenadas 445
— curvilíneas 187
— cilíndricas 191
— esféricas 191, 228
Criterio de Cauchy 24, 311
— — comparación 227, 361
— — convergencia uniforme de las integrales 312
— — Darboux 157
— — Dini 362
— — la diferencial total en una región conexa 220
— — Riemann 157
— — Weierstrass 310
Curva de Peano 133
— suave 196

Dependencia de un sistema de funciones 91
Derivada de la aplicación 68
Desarrollo asintótico 338
Desigualdad de Bessel 381, 481
— — Cauchy—Buniakovski 448

— — — Schwarz 446
— generalizada de Minkowski 172
— triangular 426
Desviación estándar 379
Determinante de Gram 469
— — Jacobi 43, 72
— — Vandermonde 548
Difeomorfismo 74
Diferencia de los elementos 422
Diferencial de la aplicación 68
— — — función 67
Divergencia 278, 281, 289

Ecuaciones de conexión (enlace) 97
Elemento del área 256
Elementos equivalentes 439, 449
— ortogonales 468
Encaje del espacio 475
— natural 464
Equivalencia entre las sucesiones 417
Espacio completo 415
— con convergencia 511
— conjugado 513
Espacio de Banach 441
— — dimensión finita 424
— — — infinita 426
— — funciones generalizadas 518, 525
— — Hausdorff 563
— — Hilbert 452, 477
— funcional 467
— lineal 422
— — complejo 422
— métrico 412
— normalizado 426
— — completo 441
— seminormalizado 426
— topológico 563
Extremo 27, 98
— condicionado 98

Factor de convergencia 324
Filtro 564, 565
— completo 566
— de Fréchet 565
— natural 565
Finura de la partición 134
Flujo del campo vectorial a través de la superficie 280, 301
Folio de Descartes 85
Forma métrica principal del espacio 251
Fórmula de cuadratura 550, 553
— — Green 203, 206, 207, 223
— — incrementos finitos de Lagrange 19
Fórmula de Lagrange 19
— — inversión 401
— — las parábolas 551
— — los rectángulos 551
— — — trapecios 551, 552
— — Ostrogradski—Gauss 285—289
— — Simpson 551, 552
— — Sojotski 520
— — Stirling 338
— — Stokes 290—295
— — Taylor 11, 12, 15, 19, 537—540
Frontera suave a trozos 210
Función absolutamente integrable 331, 345
— armónica 97
— beta 325
— característica 351
— de Heaviside 509, 522
— — Lagrange 102
— delta 506, 517
— escalonada finita 352, 497
— finita 351, 352, 497
— gamma 325
— generalizada 516, 519—524
— implícita 36
— incompleta 342
— integrable 136, 223
— localmente integrable 516
— T-periódica 349
Funcional 63, 510, 512
— lineal 63
Funciones coordenadas 52, 60

Geometría interior de la superficie 252
Gradiante de la función 248, 276
— del vector 277

Homeomorfismo 58, 76, 260
— de pegamiento 260
— — — conexo 261
Homogeneidad de la norma 426

Identificación 417, 439, 452
Igualdad de Parseval 383, 483, 484, 492, 493
Integral convergente 307, 344
— curvilínea de primera especie 193
— — — segunda especie 196
— de Darboux 153
— — Dirichlet 356, 396
— — Euler de primera especie (función beta) 325

— — — — segunda especie (función gamma) 325
— — Fourier 393
— — Laplace 404
— — Poisson 227
— — Riemann 136
— — superficie 267—275
— dependiente del parámetro 161, 301, 306
— impropia 224, 307, 330
— reiterada 161
— uniformemente convergente 308, 344
Isomorfismo 425

Jacobiano (determinante de Jacobi) 43, 72

Límite de la aplicación según un filtro 569
— — la sucesión de puntos 415, 511
— — un filtro 568
Linealidad de la transformación de Fourier 402
Líneas coordenadas 188

Matriz de Jacobi 43, 71, 91
— del operador lineal 62
Método de introducción del factor de convergencia 324
— — horquilla 547
— — las cuerdas 542
— — — tangentes (de Newton) 542, 545, 548
— — medias aritméticas 371
Medida de Jordan 118
Métrica (distancia) 413, 439
Multiíndice 18
Multiplicación escalar 445, 493
— semiescalar 445, 493
Multiplicadores de Lagrange 102

Norma 65, 426, 430, 431, 433
Normal exterior 266
— interior 266
Núcleo de Dirichlet 356
— — Fejer 371
— — la aplicación 425
Notación compleja de la integral de Fourier 400
— — — las series de Fourier 392
Nudos 554
— de interpolación 548
Números de Bernoulli 343

Operador 61
— acotado 432, 433
— continuo 514
— de Laplace 87, 223
— lineal 61, 433, 436
Orientación de la superficie 258, 264
— — — frontera 202, 206
— del borde de una superficie 265
— — contorno 202
— negativa 258
— positiva 258
Ortogonalidad 346, 468

Paralelogramo coordenado (curvilíneo) 189
Partición de la superficie 272
— del conjunto 134
Pesos 554
Plano tangente 246
Polinomio de interpolación 548, 550
— — Lagrange 549
— — Taylor 16
— trigonométrico 376
Polinomios de Legendre 470, 476
Portador de la función 351
— — — superficie 241
Potencial 276, 345
— newtoniano 345
Principio de localización 359
Producto escalar 446
— semiescalar 446
Prolongación de la función 20, 349
— — — funcional 513
Propiedad de monotonía 120
Punto autoadherente 85
— de contorno 241
— — extremo 27
— — la superficie 238, 241
— — máximo estricto 26
— — mínimo estricto 26
— — retroceso 85
— estacionario 29
— interior 241
— límite 241
— múltiple 85, 238, 241
— regular 360
— singular 77, 348
— — de la función 348

Reflexividad 561
Región elemental 286
— simplemente conexa 216, 297

Relación de equivalencia 457, 561
Representación coordenada 240
— de la superficie 237
— paramétrica 237
— vectorial 240
Rotor 278, 281, 294

Seminorma 426, 447
Serie asintótica 338
Serie convergente 444
— de Fourier 348, 362, 364, 368, 384, 389—393, 480
— — Leibniz 368
— — Stirling 343
— — Taylor 26, 537
— trigonométrica 346, 349
Símbolo de Kronecker 424
Simetría 561
Sistema cerrado 486
— completo 379, 445, 475
— linealmente dependiente 469
— — independiente 468
— ortogonal 468
— ortonormalizado 468
Subespacio 414, 423
— tendido sobre los vectores 108
Subfiltro 566
Sucesión acotada 437
— asintónica 338
— convergente 414, 436, 510, 515
— en delta 510, 519
— fundamental 415, 440
Sucesiones equivalentes 417
Suma de Darboux 144
— — Fejer 371
— — Fourier 355, 357
— — la serie 444
— integral de Riemann 135, 199
— parcial de Fourier 357
— — — n-ésimo orden 523
Superficie 237, 240
— continua 237
— dada implícitamente 244
— derivable 238, 243
— no orientable (unilateral) 264, 266
— orientable (bilateral) 262, 266
— orientada 259, 265
— suave 250
— — a trozos 262, 266

Teoremas de encaje 435
Término residual de la fórmula de Taylor 12, 15
— — — — interpolación 549
Transformación de Fourier 401, 403, 407, 504, 526
— inversa de Fourier 401, 403
Transitividad 561

Valor principal de la integral 399

A NUESTROS LECTORES:

Mir edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y de ciencia ficción. Dirijan sus opiniones a: Editorial Mir, 1 Rizhski per., 2, 129820, Moscú 1-110, GSP, URSS.